UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona

Multi-Connectivity in 5G New Radio: Configuration Algorithms and Performance Evaluation

A Master's Thesis

Submitted to the Faculty of the

Escola Tècnica d'Enginyeria de Telecomunicació de Barcelona

Universitat Politècnica de Catalunya

by

Melisa López Lechuga

In partial fulfilment

of the requirements for the degree of

MASTER IN TELECOMMUNICATIONS ENGINEERING

Advisors: Daniela Laselva

Nokia Bell Labs

Klaus Pedersen

Nokia Bell Labs & Aalborg University

Nurul Huda Mahmood

Aalborg University

Co-advisor: Jordi Pérez-Romero

UPC

Aalborg, Denmark, 2018

Title of the thesis: Multi-Connectivity in 5G New Radio: Configuration Algorithms and Performance Evaluation

Author: Melisa López Lechuga

Advisor: Daniela Laselva, Nurul Huda Mahmood, Klaus Pedersen

Co-advisor at UPC: Jordi Pérez-Romero

Abstract

The 5th Generation (5G) New Radio (NR) air interface is expected to be the foundation of very heterogeneous networks serving a wide range of use cases, including Ultra-Reliable Low Latency Communications (URLLC) services. In URLLC, small data packets must be correctly transmitted and received in a short time with high reliability (up to 1 ms latency with a success probability of 99.999%). Different options are being considered to meet this challenging design target. One such considered solution is data duplication through dual connectivity, where the same packet is independently transmitted through two different nodes.

This project studies the functionality of data duplication at PDCP level for dual connectivity through system-level simulations, where each packet copy is sent through the two links to which a certain UE is connected. The studied scenario is a heterogeneous network of 21 macro cells with a cluster of 4 pico cells per macro cell area. The scenario is first optimized for the single connectivity case, which supports up to 8Mbps URLLC load while meeting the URLLC requirements. When dual connectivity is enabled, in a controlled manner, in a URLLC traffic only scenario, it is shown that dual connectivity does not provide any gain due to the low interference conditions. As second step, the benefit of DC is studied when the URLLC traffic coexist with full buffer background eMBB traffic. Results show that latency gain can be obtained by dual connectivity, however the sensitivity of this gain on the scenario conditions is quite high. Finally, an optimization is added, in which if a packet sent through one of the links is successfully received at the UE, the transmission of its copy on the other link is cancelled (i.e. the packet is discarded at the network side). This optimization results in a performance improvement in terms of the latency especially at high load because it avoids buffering delay.

Acknowledgements

First, I would like to express my gratitude to my supervisors Nurul Huda Mahmood, Daniela Laselva and Klaus Pedersen for their support and guidance. It was a great opportunity to come to Aalborg and learn from them. Their encouragement and technical knowledge have been key for me to finish this master thesis. I have learnt and improved a lot thanks to their advices.

Additionally, I am very thankful to Mohammed, for the good discussions and the funny moments when gain seemed to be an impossible goal. I am also very grateful to Jens Steiner, for all the patience and help provided with the software implementation. Thanks also to Guillermo Pocovi and Niko Kolehmeinen, for their contributions and guidance in some aspects of the project. I would also like to thank Mads Brix for his assistance with all the IT issues and the rest of the people in Nokia Bell Labs for a nice and friendly working environment. I really enjoyed working here for the last 8 months.

Also, thank you to Jordi Pérez, my co-advisor at UPC, for the quick responses whenever I needed something and for his suggestions on this report.

On the personal side, many thanks to all the friends in Spain that were close to me even in the distance. I would like to thank also the friends that I have made here in Aalborg for making a great experience from these months. Specially, Roberto and Maria, for being incredibly funny, and such some honest and good friends.

I want to express my biggest gratitude to the most influential people in my life: my family and Enric. Despite 2500km of distance, their unconditional love and support has been essential these months. I owe them all my successes.

Revision history and approval record

Revision	Date	Purpose
0	02/10/2017	Document creation
1	24/04/2018	Document revision
2	02/05/2018	Document approval

Written by:		Reviewed and approved by:	
Date	29/04/2018	Date	02/05/2018
Name	Melisa López Lechuga	Name	Nurul Huda Mahmood Daniela Laselva Klaus Pedersen Jordi Pérez-Romero
Position	Project Author	Position	Project Supervisors

Table of contents

List of Figures

List of Tables

1. Introduction

1.1. Statement of Purpose

The next generation mobile network, also known as the 5G network, will be a heterogeneous network serving several different use cases and comprising not only the evolved existing systems but also the new radio system, based on new interface New Radio (NR) and spanning over the whole 5G NR frequency spectrum. Multi-Connectivity (MC), where a User Equipment (UE) is aggregating the radio resources from multiple cells/base stations, provides an efficient way to fulfil the 5G requirements of high data rate and ultra-reliability by leveraging different layers of the mobile network and deployments. There have been efforts in 3GPP to specify the multi-connectivity solutions like dual connectivity or carrier aggregation. This project will study the performance of MC in 5G New Radio and investigate efficient MC scheduling and configuration algorithms. The study will be focused on Ultra-Reliable Low Latency Communications use case. Starting with the dual connectivity (UE connected only to two different nodes), the configuration and data duplication control that provide the best performance for this service is studied through dynamic system level analysis. The system-level simulator used has a high level of realism on the NR technology characteristics and performance. Further sections in this report make use of specific characteristics that were already available on the simulator as well as new features implemented as a contribution from this project.

1.2. Methods and Procedures

This section includes a summary of the main research questions as well as different tables and 3GPP specifications that need to be considered to carry this project out.

1.2.1. Main Research Questions

The project can basically be divided into two different parts: Multi-Connectivity Mode Configuration and Data Duplication Configuration for URLLC.

The former should solve questions related to the decision of which nodes should take part of the Multi-Connectivity mode for a specific UE. The research in this case would be focused on the following questions:

- *What is the best URLLC performance that can be achieved with single connectivity? What is the best configuration for the baseline?*
- *What is the optimal criteria for the Master gNode-B (MgNB) to consider a node for Multi-Connectivity? That is, which is the optimal Secondary Cell (SCell) selection criteria option?*
- *For a given UE, is the MgNB sufficient to meet the service requirements? Is it optimal to apply MC? For which nodes?*

In addition, data duplication configuration leads to two basic questions:

- *How to configure PDCP Protocol Data Units (PDU) Duplication to avoid unnecessary duplication? How to optimize the duplication configuration algorithm?*
    - *Which is the best way to perform PDCP PDUs duplication?*
    - *How long should the duplication configuration last?*

These questions and the different options to study them will be further explained in Section 3.1.

1.2.2. Requirements and specifications

The aim of this project is to study the performance of multi-connectivity for URLLC in a realistic multi-cell multi-user scenario. The main idea is to duplicate data packets (Packet Data Convergence Protocol – PDCP – Protocol Data Units – PDUs) to ensure the arrival of the packets with a 99,999% of reliability, which is one part of the requirements of URLLC. At the same time, packets should arrive with a latency below 1ms.

Using as a baseline the same scenario when the UEs are in single connectivity mode (i.e. having a single cell), which should be optimized as much as possible, the performance will be compared for the different configurations proposed further in this report.

Firstly, the study is based on the dual connectivity mode, which means that the UE will be served at most by two gNBs. The study will be done in a heterogeneous network, in which the macro cell and small cell layers are at different frequencies. Therefore, data duplication for URLLC will be studied in an inter-frequency dual connectivity scenario.

The project is mainly carried out through system level simulations of a multi-user multi-cellular system. The simulator used is called FREAC. It is developed by Nokia Bell Labs and allows LTE and NR network performance evaluation under high realistic conditions. However, data duplication for dual connectivity feature was not available in the simulator and some effort had to be put on including it. Further description on the different parameters for the performed simulations, the simulator itself and the new implementations, can be found in coming sections. In addition, analytical tools, such as Matlab, will be utilized during the algorithm development phase to post-process obtained data in the simulations.

1.3. Research Plan

The research plan tabulated below maps the project into five research activities with distinct objectives and associates them to clearly defined research tasks.

Table 1. Research Plan

Research Activities (RA) and Research Objectives (RO)	Tasks (T)
RA1: Background study and Simulator acquaintance	T1.1: Identification of relevant scenarios of definition of the study cases.
RO1.1: In depth study of State of the art. RO1.2: To familiarize with the FREAC simulator.	T1.2: Identification and setting up of simulation environment.
RA2: Baseline Optimization RO2.1: Obtain best single connectivity performance in the research scenario.	T2.1: Optimization of the performance of the baseline scenario.
RA3: MC Performance Evaluation	

RO2.1: To evaluate the performance of MC considering the identified scenarios.	T2.1: In depth performance evaluation targeting URLLC services. (E.g. Understanding the reliability KPI variation when increasing the traffic load compared to no duplication as baseline. The event threshold values to trigger data duplication could be varied too.)
RO2.2: To gain insights into different factors affecting the performance of MC.	
RA4: MC configuration algorithm development	T3.1: Proposing MC configuration algorithms
RO3.1: To propose dynamic autonomous MC configuration algorithms for URLLC and eMBB services.	T3.2: Performing computer simulations to validate the proposed algorithms under standard 5G NR scenarios.
RA5: Data Duplication algorithm optimizing	T3.1: Proposing different PDCP PDUs duplication algorithms.
RO4.1: To propose the optimal way to duplicate PDCP PDUs.	T3.2: Performing computer simulations to test and validate the best duplication algorithm.

1.4. Problems and Deviations

Learning the state-of-the-art, such as dual connectivity modes and algorithms, the different features and actions that take place at each level of the protocol stack, took a great part of the time at the very first phase of the project. However, the most restricting factor in the project development was the need for understanding the system level simulator. Facing a huge C++ code implementation, developed by several persons, understanding how it works and how are the algorithms designed is very time consuming and not easy at all. Not only for the implementation of new features, but also for understanding how the statistics were collected, or why the results did not match with expectation, a deep understanding of some parts of the code needed to be done. Deep debugging of the code needed to be done for some cases, which was very time-consuming.

Furthermore, all the characteristics of the chosen scenario require a lot of configurations and processes in the system level simulator that make the simulations take a long time to finish. To have trustable results, simulations for the last part of this report included 5 million packet samples to generate the CCDF and obtain a proper accuracy at the distribution level of 10^{-5} requested by URLLC traffic. For reaching that, the simulations are rather time consuming.

2. State of the art of the technology used or applied in this thesis

2.1. Wireless Technologies Evolution

When looking backwards, it seems that it approximately takes 10 years of research and standardisation to introduce a new generation of mobile communication technologies. The last three decades, these technologies seemed to have hugely evolved. The appearance of the 3rd Generation (3G) lead to the use of the mobile phones, not only for making calls and sending SMS, but also to access the Internet by introducing the term 'mobile broadband'. In 2008 there is the very first release of the 4th Generation (4G), also known as Long Term Evolution (LTE), aiming to provide higher data rates to the users and lower complexity of the systems. Release 8 was the first of a series that meant the beginning of the development of data-oriented devices.

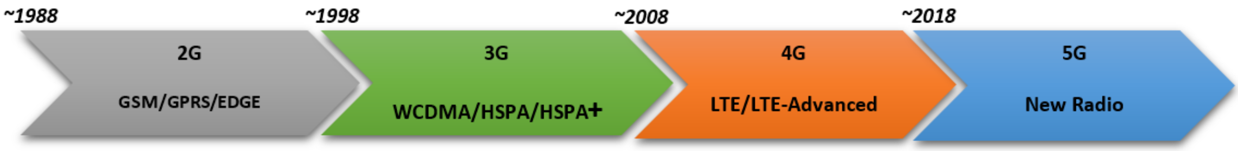

Figure 1. Timeline of 3GPP Technologies Standardization

Agreements in the development and maintenance of all these technologies should be made to motivate their utilization globally. With this need, the 3rd Generation Partnership Project appeared in 1998, which is formed by different groups of telecommunications associations (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC) and oversees the standardization of the different technologies. The 3GPP objective is to produce technical reports and specifications for each of the technologies, scoping specifically three different groups: Radio Access Network (RAN), Services & Systems Aspects (SA) and Core Network & Terminals (CT).

2.2. 5G Motivation and Requirements

The demand for better mobile broadband experiences is continuously increasing. A huge number of devices require connection to the Internet such as mobile phones, computers, tablets, smart watches, etc. Users of these kinds of devices always demand higher data rates. Moreover, several industries are going through a digital transformation. There are emerging concepts for smart vehicles, human machine interaction, sensor network, critical control of remote devices, etc. Communication procedures addressed for these new services are different from the human use case. Thus, to assure a certain quality of the service, appropriate requirements in terms of reliability, latency, throughput, scalability and energy-efficiency should be met. The current Radio Access Technologies (RATs) are not able to deal with all these requirements. Therefore, 3GPP defined the New Radio (NR) access technology (also known as the 5Th Generation or 5G), a new air interface which is supposed to handle all the capacity and performance requirements of the emerging services.

5G should be able to deal with several distinct service types such as massive Machine Type Communication (mMTC), enhanced Mobile Broadband (eMBB) or Ultra-Reliable Low Latency Communications (URLLC). Since this implies a very heterogeneous network, it suggests the need of a high flexibility, so the network can adapt to the different requirements depending on the type of service being required. The 5G NR frequencies range from low bands below 6GHz to higher bands above 30GHz (millimetre Wave - mmW).

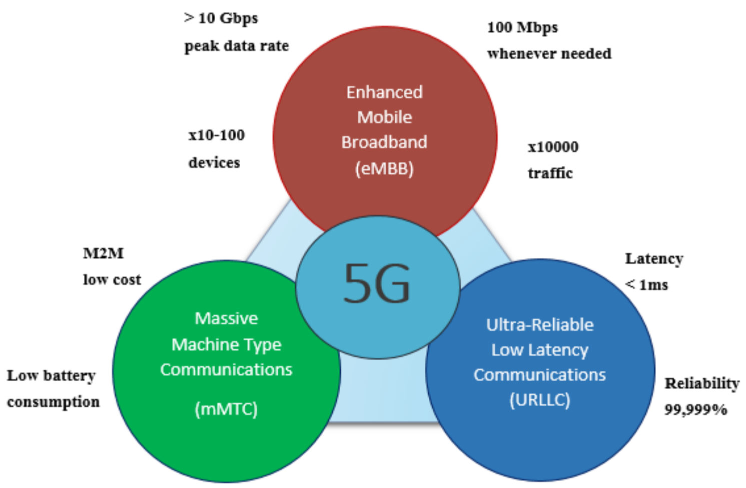

Figure 2. 5G Service Types and Requirements

The aim of mMTC services is to provide efficient connectivity for hundreds of devices per square kilometre. These devices usually send short packets and have low delay restrictions. The requirements, therefore, are low cost, low battery consumption and high connection density. Applications of this kind of service include e-Health, transport and logistics, smart agriculture, smart energy network, etc.

On the other hand, eMBB addresses human-centric use cases. It aims to improve consumer experience when accessing to multimedia content, services and data and focuses mainly on services with high bandwidth requirements.

Lastly, URLLC is intended for mission-critical links, which means that it has strict end-to-end latency and high reliability requirements. Some of its applications are vehicle-to-everything communication, drone delivery and smart manufacturing. Further explanation about these services and the requirements for the scenarios where they are used can be found in [2, 11, 12].

The European Union Commission, jointly with telecommunications operators, industry manufacturers, service providers, small and medium-sized enterprises and researches, formed the 5G Infrastructure Public Private Partnership (5G PPP). 5G PPP is also working to deliver solutions, architectures and standards input for NR. This project is partly developed in the context of ONE5G, which is one of the Phase 2 Projects of 5G PPP.

2.3. Ultra-Reliable Low Latency Communications (URLLC)

Ultra-Reliable Low Latency communications is an important service class in 5G New Radio (NR). It is intended for mission-critical communications and is the enabler for a vast set of applications. Services such as advanced energy networks, self-driving cars or intelligent industrial processes that have in common the need of very low end-to-end latency and/or high reliability. For this reason, there are several different scenarios and requirements comprising from a demand of reliability of $1\text{-}10^{-9}$ to $1\text{-}10^{-5}$, and a latency of 1ms to 10ms [6].

2.3.1. **Requirements**

In the most demanding case, URLLC general requirements according to 3GPP state that short packets (32 bytes) should be transmitted and received within a 1ms latency window with 99,999% $(1 - 10^{-5})$ reliability. However, as previously said, these requirements could be modified depending on the URLLC scenario, since there might not be the need to fulfil both at the same time. According to 3GPP agreements [21]:

- *User plane latency (L) is defined as the time it takes to successfully deliver an application layer packet/message from the radio protocol layer entry point to the radio protocol layer exit point via the radio interface in both UL and DL directions, where neither device nor base station reception is restricted by Discontinuous Reception.*
- *Reliability is defined as the success probability R of transmitting X bits within the user plane latency (L) at a certain channel quality.*

Latency would then be the period that comprises all the steps from the assignment of the transmission resources until the receiver finishes processing the data, as shown in Fig. 3. Reliability can also be defined such that the ratio of lost, erroneous or delayed messages is very low.

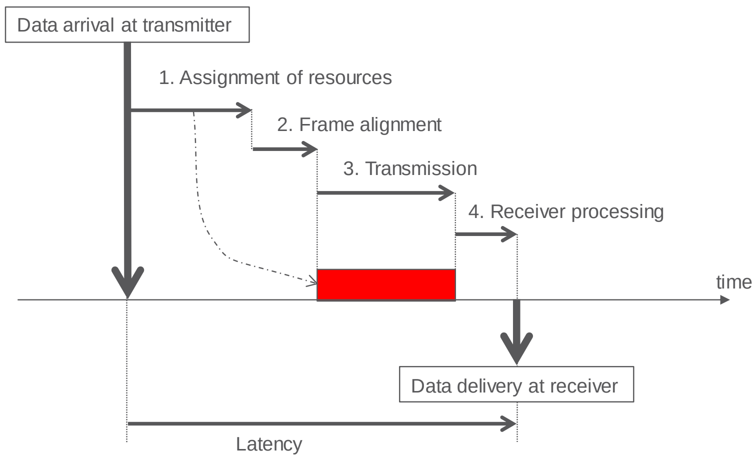

Figure 3. Latency Components in a successful reception case. Extracted from [5].

The most common way of illustrating latency and reliability is by using the Complementary Cumulative Distribution Function (CCDF) of latency. This function allows to see the probability of latency being less than or equal to a certain value. This probability can be seen as the reliability that a certain packet is received within the required latency.

In a wireless system, it is challenging to fulfil simultaneously stringent reliability, latency and throughput due to the main sources of variability, namely the traffic source and the wireless channel [2]. As a result, there is a fundamental trade-off among latency, reliability and throughput in a cellular network, i.e., for a given latency and reliability target, there is a maximum achievable rate. Fig.3 shows the latency components under ideal conditions. For that case, data has been sent right after arriving to the buffer (i.e., without queueing delay) and correctly decoded at its destination without lower layer retransmissions. There could be the case where the data arrives to the buffer and cannot instantly be scheduled, so the queueing delay will be added to overall latency, because of the required buffering time. Data can also be incorrectly delivered for several reasons and one or more retransmissions might be needed to transmit correctly a given packet, which would also increment the packet delay. The offered load on the cell also plays an important role affecting the delay, since an overload situation would increase the number of packets queued prior to transmission, and therefore, increase their queueing delay.

2.4. Multi-Connectivity

The trade-off explained in the previous section, remains when multi-node connectivity is considered. Multi-Connectivity (MC) can be utilized to improve the data rate and reduce the latency simultaneously.

MC consists on aggregating radio resources from distinct nodes to serve a single User Equipment (UE). By changing the configuration, Multi-Connectivity can provide different benefits as explained in the following. Splitting the data through the different nodes to which the UE is connected, would provide throughput enhancement for both uplink (UL) and downlink (DL). It also allows seamless mobility. For the case of data duplication, it is expected to provide ultra-reliability and latency reduction. A multi-connectivity capable user is expected to be able to support simultaneous connectivity and aggregation of resources even though they are from different technologies (LTE, Wi-Fi, etc.). It also should be able to connect to different network layers (macro and small cells) and different RAT layers.

Depending on the frequencies of the aggregated links, MC can be categorized into: inter-frequency MC, where multiple links at different frequencies are aggregated; and intra-frequency MC, targeted at boosting the performance of cell-edge users. This work will primarily focus on inter-frequency MC.

The hypothesis is that MC is a powerful way of increasing reliability. Nevertheless, most of the existing concepts and studies on MC focus on data rate boosting. The investigations related to reliability boosting (i.e. data duplication) via MC are less available (so far) and have a higher level of abstraction, see [4]. Studies regarding the optimal 5G architecture for multi-connectivity to improve reliability performances can be found in [7]. Specifically, works targeting performance aspects with URLLC services and proposing dynamic configuration algorithms are limited, which leaves room for contribution in this domain.

2.4.1. **Dual Connectivity**

Dual Connectivity (DC) was first introduced by 3GPP in Release 12 for LTE. In this technique, the UEs are configured to utilise radio resources from two different eNode-Bs (eNBs) connected via a non-ideal backhaul over the X2 interface. There is a Master eNB (MeNB), which maintains the control plane, and a Secondary eNB (SeNB). There is only one C-plane S1-MME connection per UE, Radio Resource Connection (RRC) is established only via MeNB, which also controls the SeNB connection. User plane is split between both.

In a typical scenario, as the one shown in Fig.4, the MeNB is a macro cell while the SeNB is a small cell. Dual Connectivity is a form of Multi-Connectivity.

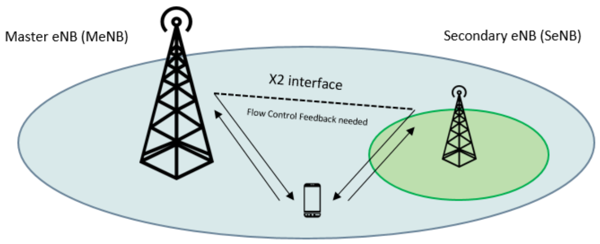

Figure 4. Typical DC Scenario

There are two different DC solutions depending on where the split of the user plane is done. It can be split in the Core Network (CN), so that MeNB and SeNB serve separate radio bearers (1A type, architecture and U-plane protocol stack available in Fig.5); or in the MeNB (3C type, architecture and U-plane protocol stack available in Fig.6). In the latter case, the data of a different radio bearer can be transmitted via both the MeNB and the SeNB, so this offers higher flexibility. However, this is at the cost of increased transport and processing capabilities in the MeNB and in the X2 interface. There is also the need of flow control feedback to avoid congestion and increase the rate control efficiency across eNBs [13,14]. The 3C architecture is the one used for the system level simulations performed in this project.

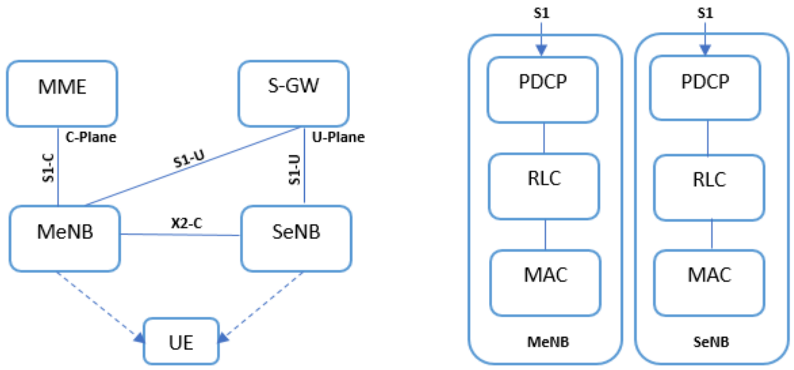

Figure 5. 1A DC Architecture and U-plane Protocol Stack

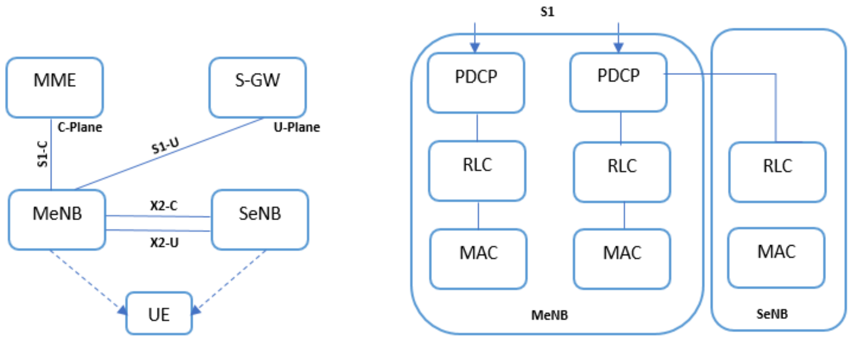

Figure 6. 3C DC Architecture and U-Plane Protocol Stack

2.4.1.1. Dual Connectivity and Carrier Aggregation

The first LTE release (Rel-8) was introduced in 2008, but it was not until March 2015, with LTE-Advanced (Rel-12) that dual connectivity appeared mainly to support small cell enhancement and increasing user throughput. It was first only introduced supporting downlink, and including a MeNB and a single SeNB. There were changes in S1 and X2 interfaces to support DC. Release 13 (2016) included support also uplink split bearer as well as several other features that improved dual connectivity performance. Is in Release 14 (2017) when it starts the standardization of 5G. There rises the idea of extending the concept of dual connectivity to multi-connectivity. Letting the terminal connect to several nodes could be a good way to meet the tight requirements in terms of latency and reliability of some of the 5G services.

Dual Connectivity can be easily confused with Carrier Aggregation (CA). This section aims to clarify the main differences between these two techniques. Carrier Aggregation consists basically of allowing a UE to transmit and receive data simultaneously on multiple component carriers. The carriers, in this case, pertain to a single eNB.

As previously explained, DC uses radio resources from multiple carriers to improve user throughput, which is also the main objective of CA. CA and DC are usually applied in different scenarios depending on whether the delay backhaul is ideal or not, respectively. Therefore, there are also some differences in their design. The main difference Is that the user plane split in CA is done at Medium Access Control (MAC) level, while in DC, it is done at PDCP level. The fact that they are different techniques does not mean that they are mutually exclusive. In fact, both can be applied to the same UE at the same time. Studies in [9] show the performance of DC can be close to the one with inter-site CA with ideal fiber-based fronthaul.

So far, DC has been utilized to boost user throughput performance by splitting data between MeNB and SeNB. However, its performance in the context of 5G new radio is not well understood. This is especially true when considering the URLLC service class for which data duplication seems suitable to increase reliability, unlike data splitting for data rate boosting. In addition, enabling MC options calls for dynamic autonomous algorithms that will configure this option, for example, by determining the number of required links and selecting dynamically how much data to transfer among the selected links to meet given performance targets.

2.4.1.2. Data Duplication in Dual Connectivity for URLLC

Why is data duplication in Dual Connectivity expected to be useful for URLLC? The URLLC requirements have already been met in previous studies for a static scenario [16]. Cases with mobility (i.e., where handovers are performed) have not been investigated yet.

This project will study how the supported offered load can be increased while meeting the URLLC requirements when using DC. Also, for the same supported load, reducing the delay would be perfect for certain kind of applications. If the UE is connected to two different links, and it receives the same packet via both, the possibilities of the packet to be successfully received within the required time will increase, as explained further in this section. The DC configuration mode and duplication algorithms need to be optimized for this target to be achieved.

Terminology and Protocol Stack

As in LTE existed the eNode-B (eNB), in NR the network element of the radio access network is referred to as next generation Node-B (gNB). For the same reason, Master and Secondary cells serving a UE in DC, will be now called MgNB and SgNB, respectively. The gNBs are interconnected with each other by means of the Xn interface.

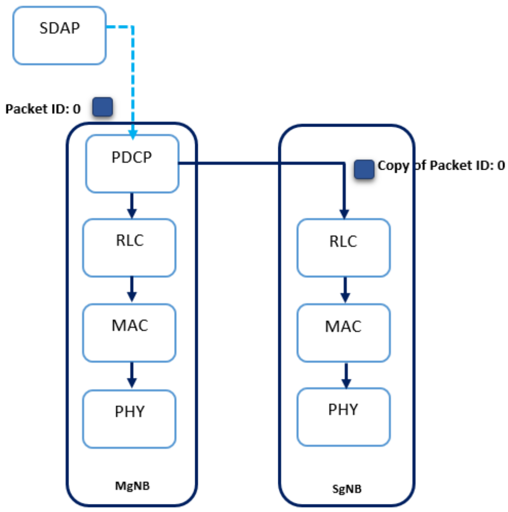

Figure 7. Data Duplication in Dual Connectivity for URLLC - Protocol Stack

Fig. 7 presents the user plane protocol stack for the dual connectivity case studied, and a simple illustration presenting how data duplication is performed. 5G New Radio Protocol Stack will generally follow the same structure than in LTE. As in LTE, it will include the physical (PHY) layer, the Medium Access Control Layer (MAC), the Radio Link Control (RLC) layer and the Packet Data Convergence

Protocol (PDCP) layer. On top of the latter, it is included a new Access Stratum (AS) layer named Service Data Adaptation Protocol (SDAP), which is in charge mainly of handling mapping between a QoS flow and a data radio bearer. The rest of the stack performs practically the same functions it was performing on LTE [20]:

- *PHY:* Transmit information over air-interface, cell selection, link adaptation, etc.
- *MAC:* Scheduling information reporting, Hybrid Automatic Repeat Request (HARQ) for error correction, etc.
- *RLC:* Transfer of upper layer PDUs, ARQ for error correction, etc.
- *PDCP:* User data transfer handling, duplication detection, **packet duplication**, header compression, etc.

The most interesting layer for this research is the PDCP layer. In the downlink direction, packets will arrive at the MgNB buffer and will automatically be duplicated at PDCP level. The duplicated packet will be sent to the SgNB over the X2 network interface. From that point, all the functions will be performed separately. This is important to consider, since it implies that scheduling and HARQ will be independent for the two gNBs. In the case of a failure in one of the links, the transmission through the other link can possibly be successfully received at the UE, thus improving the reliability.

Block Error Rate (BLER)

The BLER is a measure of how successful a data transmission is over the channel. Therefore, the BLER target is the target error rate for the transmission and it is selected based on Channel Quality Indicator (CQI) reporting. It is expected that, when the two links are transmitting the same information, the BLER target can be higher in those links since there is always a backup. That is, the BLER target condition can be more relaxed. To proof this assumption is correct, this section studies theoretically the advantages of having two links in terms of error probability. For the study, it is assumed that short TTI is used and only one retransmission can be done to fulfil URLLC requirements. These assumptions will be further explained and proved in Section 3.1. Also, next calculations do not include any queuing, processing or propagation delays.

In the case of single connectivity, having a certain error probability Pe, the probability of having a correct transmission the first time the packet is sent is:

$$P(Tx)_{SC} = 1\text{-}Pe \qquad (2.1)$$

While in the case of needing a retransmission, the probability of correctly transmitting the packet would be the probability of having an error in the first try multiplied by the probability of correct transmission in the second try. Note, here it is assumed that the feedback is without error. This is:

$$P(ReTx)_{SC} = Pe \cdot (1\text{-}Pe) \qquad (2.2)$$

Let us consider now the dual connectivity case, where there is a MgNB link with error probability Pe_1 and a SgNB link with error probability Pe_2. In this case, the probability of good transmission in the first try is given by:

$$P(Tx)_{DC} = (1\text{-}Pe_1) \cdot Pe_2 + (1\text{-}Pe_2) \cdot Pe_1 + (1\text{-}Pe_1) \cdot (1\text{-}Pe_2) = 1\text{-}Pe_1Pe_2 \qquad (2.3)$$

The case of needing a retransmission in dual connectivity, would come after the failure of both links in transmitting correctly in the first attempt. Therefore, the probability of correct transmission

$$P(ReTx)_{DC} = Pe_1Pe_2 \cdot (1\text{-}Pe_1Pe_2) \qquad (2.4)$$

It can be appreciated when comparing these probabilities that the fact of having two different links to transmit the same information allows a more relaxed BLER target (i.e., higher) value in each of the links, since there is a backup link in case of failure. When comparing Eq. 2.1 and Eq. 2.3, for example, for the same transmission probability Pe_1 and Pe_2 can be square root of Pe. There is a quadratic relation. This can also be observed in Fig. 8, which assumes the same probability for Pe, Pe_1 and Pe_2. The right plot of this figure clearly shows how the requirements of URLLC can be easier fulfilled with dual connectivity in an ideal scenario.

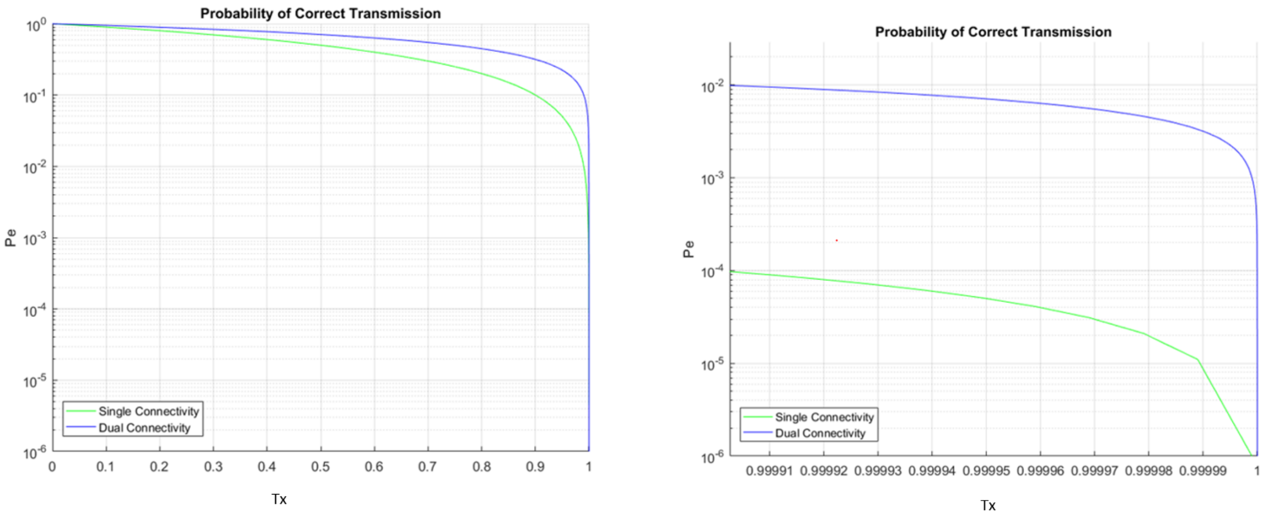

Figure 8. Transmission Probability comparison between SC and DC on the left and zoom of the same figure on the right.

There is also a **cost** on using data duplication. Since the same packet is being sent through two different base stations, and the double resources are being used for that, the spectral efficiency of the system will noticeably decrease. It will be seen that some kind of intelligence needs to be applied in duplication in order to minimize the extra resource utilization.

In addition, packet duplication increases considerably traffic in the network and, thus, interference. As mentioned above, the BLER target can be more relaxed in dual connectivity due to the existence of a backup link. However, the increasing interference will increase also the actual experienced BLER, which in turn affects the performance of the system.

Several factors affect the performance of the system. All of them need to be considered to ensure results are as much realistic as possible. For this reason, a dynamic system-level simulator is used for the evaluation.

3. Methodology / Project Development

3.1. System Level Simulator

To better understand this report, it is important to have an idea about how the adopted system level simulator works. This section briefly explains the main features of the simulator, and through the rest of the report the sub-sections named “Contribution to FREAC”, will contain a brief explanation of the author contributions to the system level simulator.

As previously stated, this simulator has been implemented by many collaborators following the LTE and NR specifications. The simulator has symbol resolution, which means that one step in FREAC corresponds to one OFDM symbol. It models PHY, MAC, RLC and PDCP layers functionalities. It includes modelling of most of the RRM functionalities such as packet scheduling (proportional fair scheduling for what it concerns this report), packet duplication and re-ordering (at PDCP level), Automatic Repeat Request (ARQ) at RLC, Hybrid Automatic Repeat request (HARQ) and link adaptation at MAC level. As recommended in 3GPP specifications, the scheduler decision for distributing the available Physical Resource Blocks (PRB) is done at every Transmission Time Interval (TTI). The user equipment reports the CQI, which allows the link adaptation functionality to select, for the set of chosen PRBs, a proper Modulation and Coding Scheme (MCS) to fulfil a certain BLER target. MCS selection is improved by using Outer Loop Link Adaptation (OLLA) mechanisms, that use the information received in ACK/NACK messages to finetune the MCS selection. The available MCS in the simulator include QPSK (1/5 to 3/4), 16QAM (2/5 to 5/6), and 64QAM (3/5 to 9/10).

To determine if the packet has been successfully decoded a link-to-system-level mapping is used, and it considers the per-subcarrier SINR of each scheduled UE. If there is an error in the transmission, HARQ is used to retransmit the packet with ideal chase combining at the receiver end. Short TTI configuration (1 TTI = 0.1428 ms) is being used for all the simulations (further explanations about numerology being used, will be given in Section 3.2.). HARQ RTT will scale linearly with TTI length. That is, it is reduced to 4 TTIs. This is done to allow at least one retransmission in case it is needed while keeping the latency below 1 ms. The components of the HARQ procedure used are presented in Fig. 9. Reducing the number of necessary TTIs for a HARQ retransmission implies improving processing capabilities at the user equipment and at the gNB.

This consideration allows performing one retransmission, which will take at most 6 TTIs (could be less than 6 depending on the packet arrival time within one TTI). Considering 6 TTIs * 0.143 ms is already 0.86 ms, one more retransmission will mean exceeding the 1 ms latency requirement.

The number of URLLC UEs is fixed during the whole simulation time and the UEs are uniformly distributed across the network. Downlink traffic only is simulated using the FTP Model 3 (FTP3), which means that small data packets, of 50 and 200 bytes, will be sent from the base stations to the UEs. Full-buffer eMBB downlink traffic is further added to stabilize the interference conditions in the later scenario.

The UE receiver type is Minimum Mean Square Error (MMSE) – Interference Rejection Combining (IRC) and links are modelled with a 2x2 closed-loop single-user MIMO.

A configuration file in the FREAC simulator allows configuring most of these parameters and having a detailed control on the simulation. However, some parameters needed to be added for the developments of this project.

3.1.1. Delay Statistics

Since delay of the FTP3 packet will be used as Key Performance Indicator (KPI), delay statistics are going to play the most important role when evaluating the results, this section explains how these statistics are collected.

The statistics used express the delay per finished FTP3 packet. The CCDF is used to plot this delay, and one sample represents the delay of one finished FTP3 traffic model packet. The delay is calculated from packet arrival to the base station buffer, until it is correctly received in the UE and reported to the BS. That is, the statistics also considers the acknowledgement (ACK) time. Since the simulator has a symbol resolution, and the length of the TTI is set to 2 OFDM symbols, packets can arrive at the beginning of the TTI (Fig. 9, option 1), in the middle of the TTI (option 2), or at the end of it (option 3). For options 1 and 3, the packet cannot be immediately scheduled, and it waits in the buffer until the next TTI. This is referred to as frame alignment time.

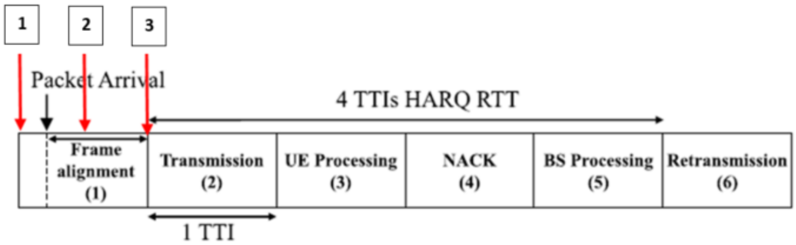

Figure 9. Diagram of HARQ operation with 4 TTIs RTT. Packet Arrival Options. Main figure extracted from [15].

For this reason, in a correct transmission, the delay is at least of one TTI and never higher than 2 TTIs. In the case of an unsuccessful transmission, HARQ process will be used to retransmit the packet. This process implies the addition of 4 TTIs to the delay. It includes the frame alignment time (if there is) and the first try of transmission time; plus, the UE processing, the Non-Acknowledgement (NACK), the BS processing, and the second try transmission times. Therefore, for the case of one retransmission, the delay is at least of 5 TTIs and never higher than 6 TTIs, as shown in Fig.10.

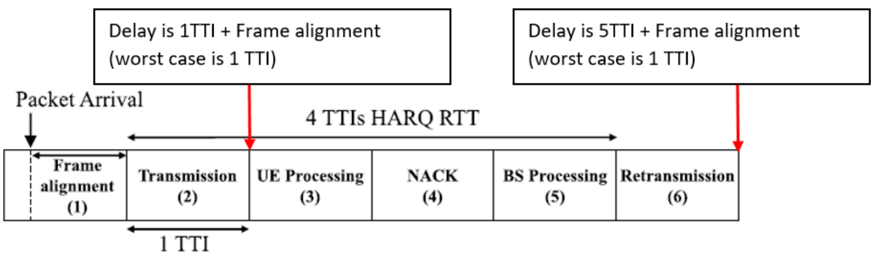

Figure 10.Diagram of HARQ operation with 4 TTIs RTT. Delay Options. Main figure extracted from [15].

3.2. Scenario

The chosen scenario is 3GPP generic defined as 2a scenario in [8]. It consists on a 7 macro sites scenario, with 3 sectors per macro and 4 picocells cluster per macro sector, as shown in Fig. 11. Separate frequencies are used for the deployment of the macro and small cells, so it is an inter-frequency scenario. Minimum distances between cells follow the specifications as well in [8] and an ideal backhaul is assumed between gNBs. Simulations are performed in a static scenario. There is no mobility of UEs, meaning that the serving cells will remain the same through the entire simulation time. However, there are still a lot of time-variant processes such as fast fading, that can cause the interferences to vary. Channel and UEs conditions might be affected by this kind of processes.

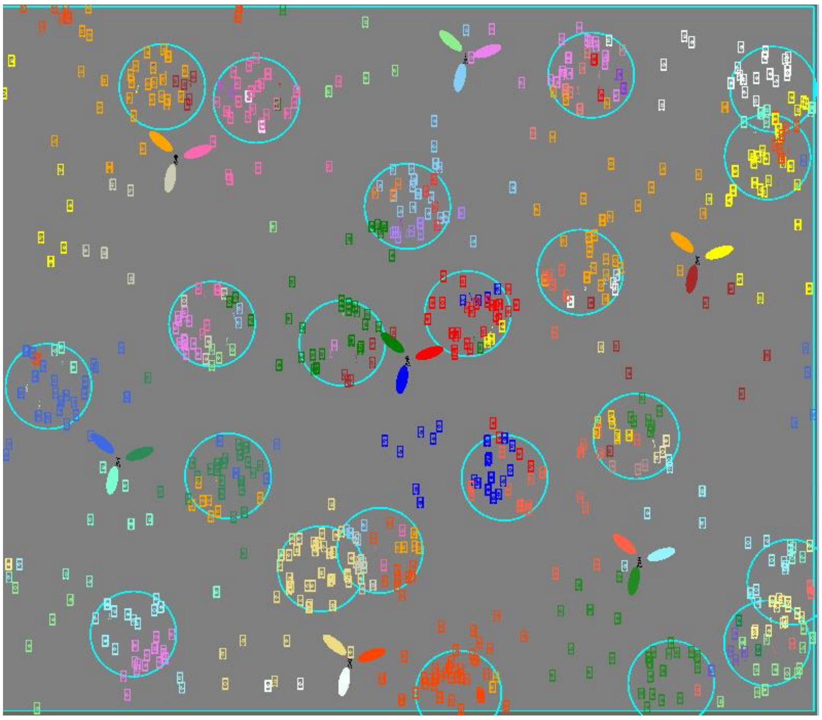

Figure 11. Research Scenario

3.2.1. **General Parameters**

The basic simulation assumptions are shown in Table 2. More detailed parameter values are presented in Table 6 of Annex A. The focus is on downlink evaluation only. There is a 15kHz spacing between subcarriers and Orthogonal Frequency Division Multiple Access (OFDMA) is used to dynamically multiplex users on a shared channel. A short TTI of 2 OFDM symbols is used (0.143 ms). In the frequency domain, a Physical Resource Block (PRB) resolution of 12 subcarriers can be used to multiplex UEs. Asynchronous Hybrid Automatic Repeat Request (HARQ) is used. A BLER target of 1% is set for the baseline simulations.

Table 2. Scenario General Parameters

Parameter	Macro Layer – NR	Pico Layer - NR
Layout	7 sites, 21 cells, wrap around	4 pico cells cluster per macro cell
Inter-BS distance	500 m	Cluster
Carrier Frequency	2GHz	3.5GHz
Bandwidth	10 MHz	
BS Power	46dBm	30dBm
Pathloss Model	3D-Uma	3D-UMi
Antenna Height	32m	10m

To establish a low, medium and high offered load supported by each of the macro cells in the baseline scenario, several simulations for different offered LLC loads were performed. The offered load per macro cell is calculated following Eq. 3.1, where λ is the arrival rate, B is the payload size in bits and N is the average number of LLC UEs per macro cell. L is given then in bits per second.

$$L = \lambda \cdot B \cdot N \tag{3.1}$$

3.3. Baseline Optimization

To ensure the performance of the configuration algorithm designed for data duplication in dual connectivity is good enough, this study aims to compare it with the most optimized single connectivity scenario. Furthermore, previous studies have been done for the single connectivity case in macro cell only scenario and for 2D channel model, which does not capture the elevation channel characteristics. For this case, a new scenario is studied for the 3D channel model (which is a bit more challenging than the 2D), and it is shown that performance improves with respect to the previous study [16]. As expected, supported load meeting the URLLC requirement increases for the studied scenario.

Using the same scenario shown in Fig. 11, URLLC traffic and short TTI, Monte Carlo simulations were performed to find out which is the best cell selection criteria and Cell Range Extension (CRE) to balance the load in the network. The same procedures were followed to decide the ranges for low, medium and high offered load of this scenario.

3.3.1. Cell Selection Criteria

In the process of cell selection, the UE looks for the most suitable cell to connect to within all the candidate cells in its surroundings. The UE can select the best cell based on two different Measurement Reports as configured by the network:

- Reference Signal Received Power (RSRP)
- Reference Signal Received Quality (RSRQ)

The RSRP is defined in [10] as the "*linear average over the power contributions (in Watts) of the resource elements that carry cell-specific reference signals within the considered measurement frequency bandwidth*". That is, a value stored in certain reference signals only, that informs the UE of the received power from the corresponding base station. On the other hand, the RSRQ is defined as well in [10] as "*the ratio N×RSRP / carrier RSSI, where N is the number of RB's of the carrier RSSI measurement bandwidth".* RSSI stands for Reference Signal Strength Indicator and it measures the quality in terms of strength of the signal received by the UE from a certain base station. It carries information about the total wideband received power. This means that in the case of RSRQ, noise and interferences are considered.

For the serving cell selection, the UE selects the one with the highest RSRP from its sets of measurements. The users are dropped at the beginning of the simulation, when it is all being configured. At that moment, there is no interference generated. For simplicity, also no UEs mobility is considered in the simulated scenario, so there would not be cell reselection for handovers. Therefore, using RSRQ as cell selection criteria would make no sense. Therefore, although system level simulations for the RSRQ case are presented in Section 3.4.1.2, RSRP is finally used as cell selection criteria. For the same reason, theoretical simulations are done for the RSRP case only.

The next sections show the performed simulations and reasonings to choose the appropriate cell selection criteria, as well as the necessary Cell Range Extension (CRE) to have balanced load in the network.

3.3.1.1. Theoretical Analysis

In this kind of scenarios, it is very typical to use a positive cell selection offset to the received power of the small cell to extend its coverage due to the transmit power unbalance. The UE would connect to a cell that is actually weaker than the strongest cell detected. This is called Cell Range Extension (CRE) and is typically used as a form of inter-layer load balancing between the different frequency layers. It allows offloading more UEs from the macro cell to the small cells, alleviating or resolving in this way the load bottleneck problems.

To determine the appropriate values for Cell Range Extension in the scenario of interest, a Matlab script was developed, which models the RSRP received in the UE from both the macro cell and the small cell. The difference between both measurements indicated the necessary range extension required at each position of the UE within 300m (which is assumed to be approximately the coverage of the macro cell) to connect to the small cell rather than to the macro cell. Since the small cells are randomly placed in the scenario following the specifications in [8], they can be at different distances from the macro cell. Therefore, three different cases where studied:

Case 1)

In this case the small cell was placed 55m away from the macro cell as shown in Fig. 12, which is the minimum possible distance between them.

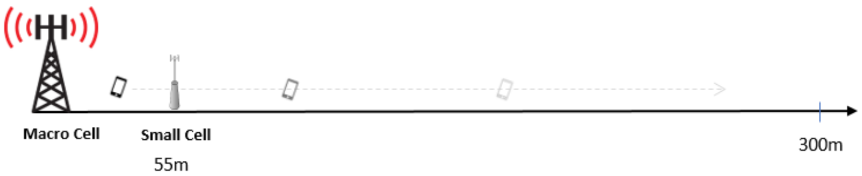

Figure 12. Theoretical Simulation Scenario for 55m distance between Macro and Small cells

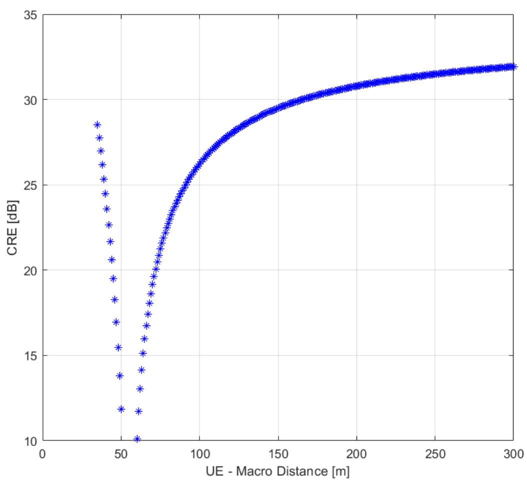

Figure 13. CRE values for 55m distance between Macro and Small cells

Case 2)

For the second case it is chosen approximately the medium distance between the minimum and the maximum, as shown in Fig.14.

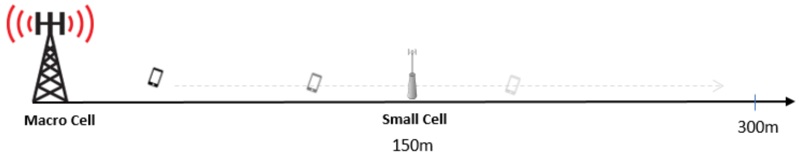

Figure 14. Theoretical Simulation Scenario for 150m distance between Macro and Small cells

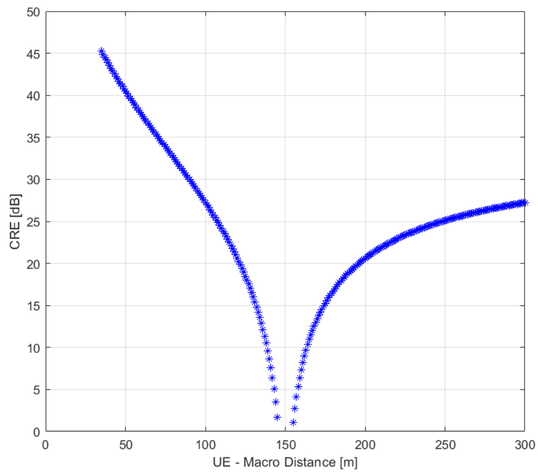

Figure 15. CRE values for 150m distance between Macro and Small cells

Case 3)

For the latest case, as shown in Fig.16, 300m distance is chosen between the two cells, which is assumed to be the maximum possible distance between them.

Figure 16. Theoretical Simulation Scenario for 300m distance between Macro and Small cells

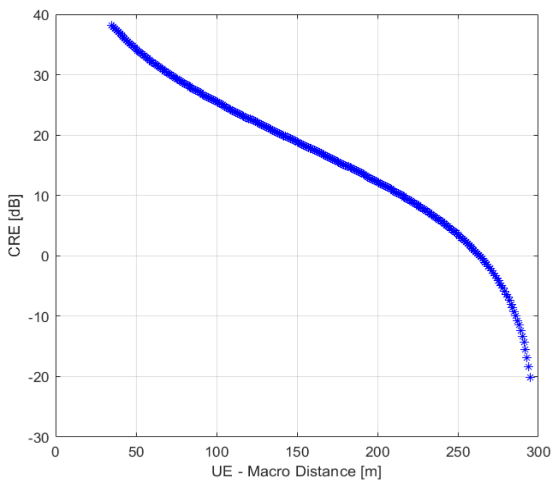

Figure 17. CRE values for 300m distance between Macro and Small cells

It is clearly appreciated in Figs. 13, 15 and 17 that as the UE approaches the small cell the necessary CRE decreases, as expected. Theoretical analysis has been done for all the possible positions of the UE. That is, assuming that a UE can be placed from 50m to 300m away from the macro cell. Therefore, as the UE approaches the macro cell, the value of the necessary CRE is unreasonably high, indicating that it would be better for the UE to connect to the macro cell. Concluding, the results for each of the three cases, show that reasonably the applicable CRE value to bias the RSRP measurements should be between 10 and 25dB.

3.3.1.2. System Level Simulations

As mentioned, system level simulations are performed using the FREAC simulator. The total number of UEs in the network is 630, which means that there should be on average 30 UEs per macro cell. As the UE dropping proportion is 1/3 to the macro cell and 2/3 to the small cell, the load balance will be achieved with 10 UEs connected to the macro cell and 20 UEs connected to the pico cells cluster (5 UEs per pico cell). Several Monte Carlo simulations are performed to adjust the CRE values range for both RSRQ and RSRP cases to reach these load conditions. It can be observed in the figures below, that for the RSRQ case (Figs. 18 and 19) the range is 0.5dB to 2 dB, which is lower than for RSRP case (Figs. 20 and 21) in which the range is 2dB to 10dB.

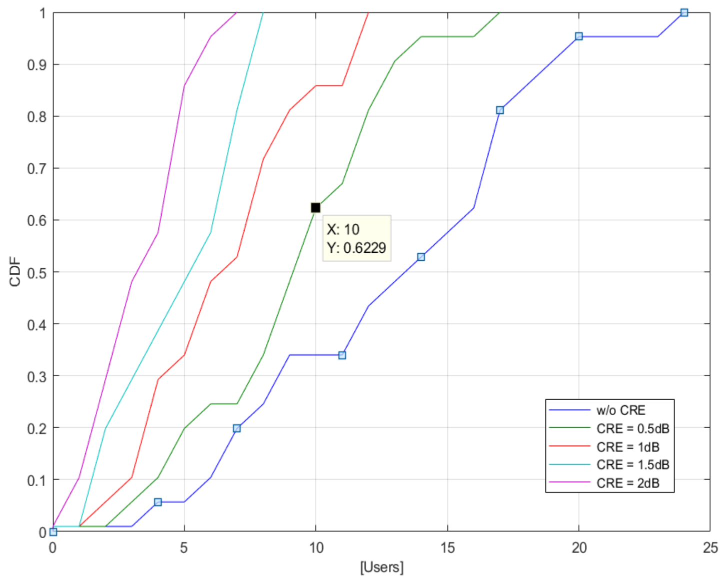

Figure 18. Macro Cell Connected Users per TTI - RSRQ Cell Selection Criteria for different CRE values

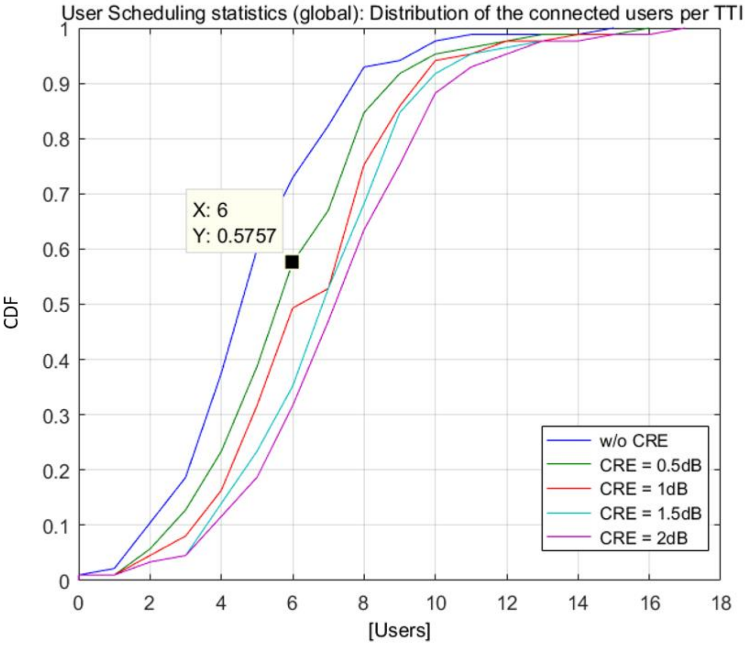

Figure 19. Small Cell Connected Users - RSRQ Cell Selection Criteria for different CRE values

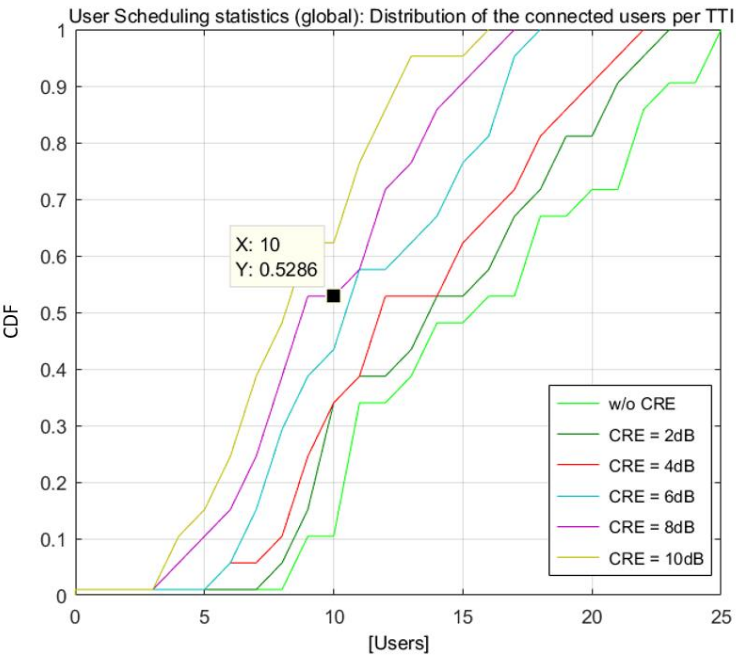

Figure 20. Macro Cell Connected Users – Slow Faded RSRP Cell Selection Criteria for different CRE values

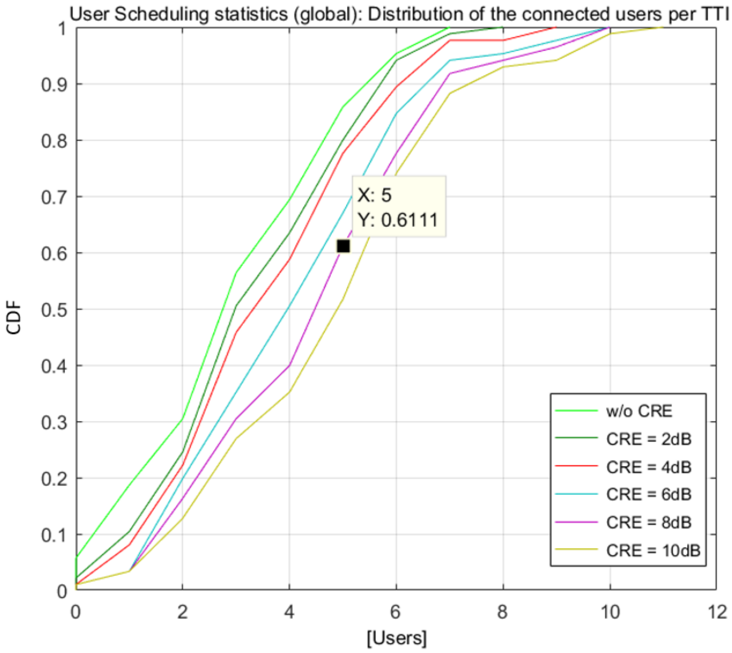

Figure 21. Small Cell Connected Users - Slow Faded RSRP Cell Selection Criteria for different CRE values

As explained in previous sections, RSRP cell selection criteria is chosen and based on these results, the CRE will be selected equal to 8dB to carry different load simulations out.

3.3.2. Cell Load

Once the CRE for achieving a load balance in terms of connected UEs across the network layers was decided, the supported offered load while meeting URLLC requirements is investigated next. Eq. 3.1 parameters were replaced with the values according to Table 3 to obtain the different loads. These values were chosen following the results in [16,17] for single connectivity in different scenarios.

Table 3. Cell Load Parameters

Offered Load per Macro Cell	2Mbps	4Mbps	6Mbps	8Mbps
λ $(Packets/s)$	41.66	83.33	125	166.66
B	200B			
N	30			

First simulations were performed with CRE = 8 dB, and, following the results on [16,17], the scenario was expected to support up to 8Mbps still meeting the requirements. As appreciated in Fig.22, with CRE= 8 dB the requirements were not fulfilled for 8Mbps load. Considering the number of macro cells (21), the number of small cells (84), how the UE are dropped (1/3 to the macro cells and 2/3 to the small cells), and the fact that both network layers have the same bandwidth, the macro seemed to be the bottleneck.

Fig.23 shows the delay per FTP3 packet per macro and small cells under the different load conditions. It is clearly observed that the macro cell is the one causing the higher delays.

Due to this fact, it was decided to increase the CRE to steer more traffic to the small cells. After various simulations, a CRE of 15 dB was chosen. It can be seen in Fig. 24 that even the 8Mbps offered load per macro simulation fulfils the URLLC requirements. Also per layer statistics for CRE = 15 dB in Fig.25 show that there is a load balance between both layers.

Different parts can be distinguished in all the CCDFs of the URLLC latency shown below. The upper part of the distribution, comprised between 10^0 and 10^{-3}, corresponds to the delay of those payloads that have been immediately scheduled and successfully transmitted with the first try. The queueing delay generated at the buffers increases with the load conditions. Then it can be observed the delay caused by the HARQ retransmissions.

Table 4 states the values decided for the baseline scenario.

Table 4. Baseline Set of Loads

Parameter	Value
Load per macro cell area	Very Low: 2Mbps
	Low: 4Mbps
	Medium: 6Mbps
	High: 8Mbps

Contribution to FREAC

To proof the bottleneck effect, some parts of the FREAC code where modified to dump the delay per network layer in a logging file. A Matlab script was then generated to obtain per layer delay statistics.

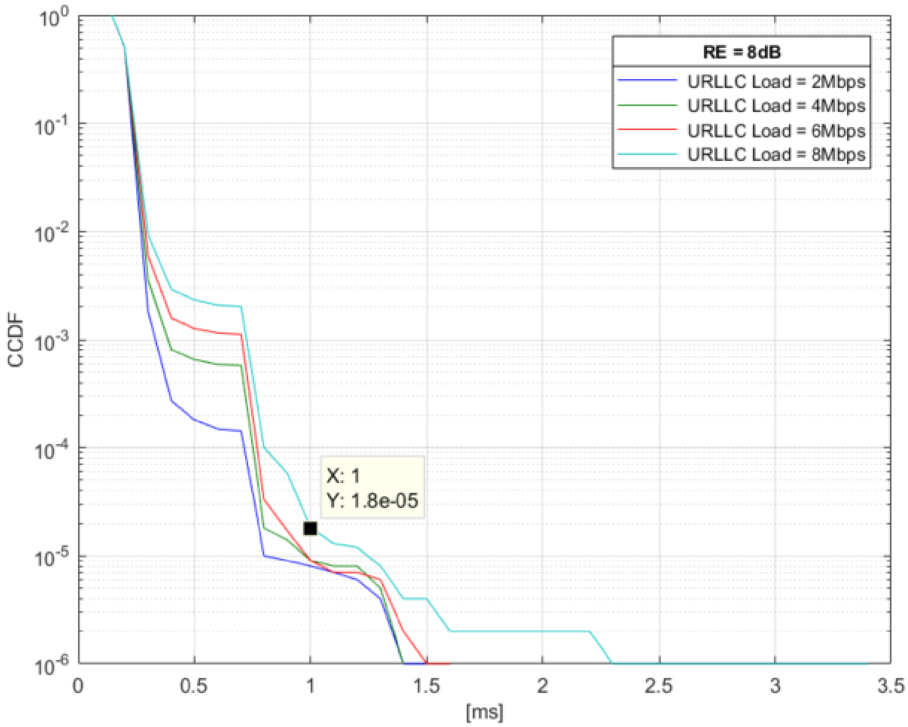

Figure 22. Delay Per FTP3 Packet for different loads with CRE = 8dB

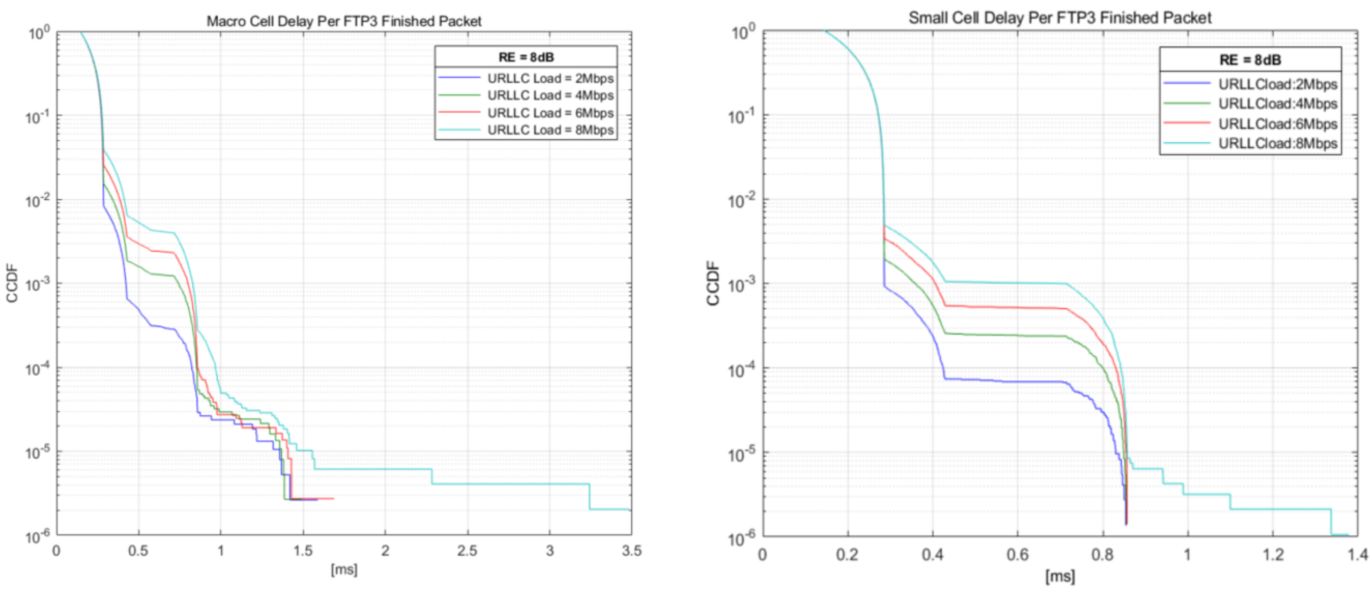

Figure 23. Per Layer Delay Per FTP3 Packet for different loads with CRE = 8dB

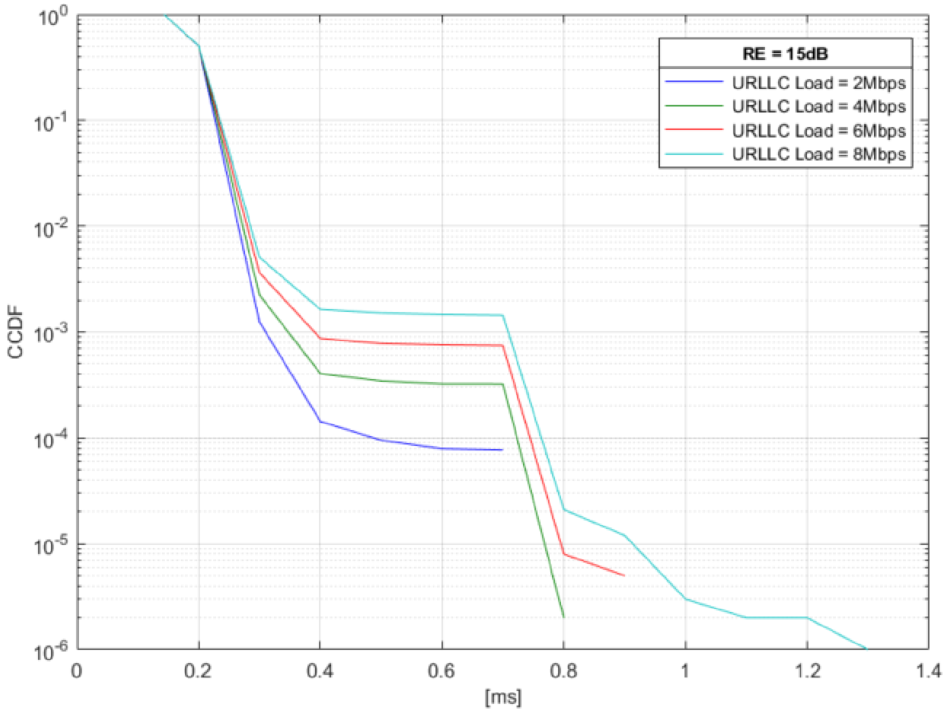

Figure 24. Delay Per FTP3 Packet for different loads with CRE = 15dB

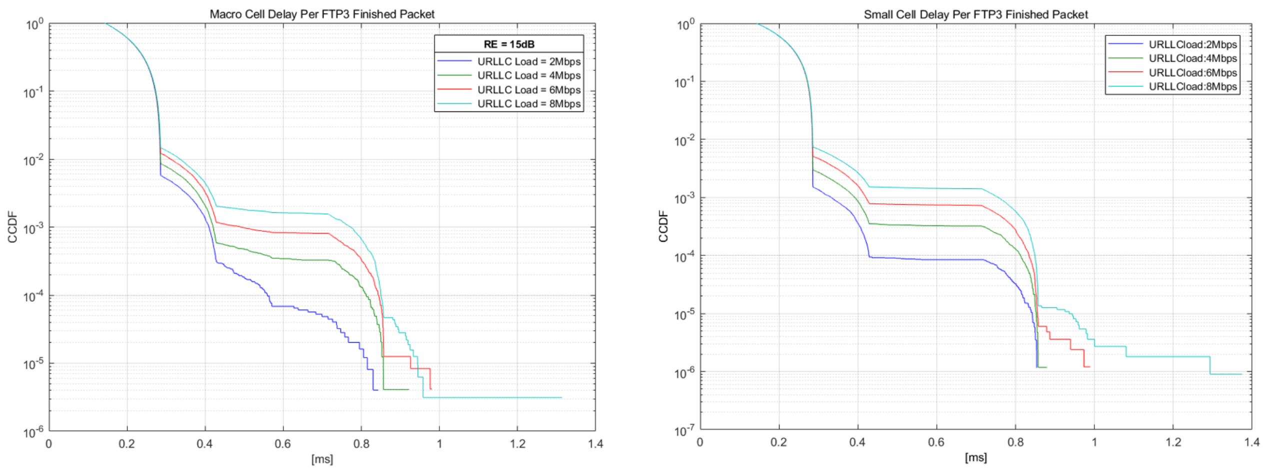

Figure 25. Per Layer Delay Per FTP3 Packet for different loads with CRE = 15dB

When comparing to theoretical simulations shown in Section 3.3.1, these results seem to be reasonable. Theoretical simulations are done for three different location cases, and samples are averaged. On the other hand, system simulations are closer to the real case, where picocells are mainly placed at macro cell edge to improve the coverage, and therefore, there are more samples for that case. This could be the reason for the mismatch in the results.

3.3.3. **Later Modifications in the Baseline Scenario: Packet Size**

Next sections show performed simulations for different BLER targets and load. At low load, low BLER target (1%) is expected to provide the better latency performance than higher BLER target (10%); while at high load, it is expected to be the opposite. If there is high load in the system, it is possible to use more aggressive transmissions. For that case, there is a delay caused by the increment of retransmissions, but it is still lower than the delay caused by the queueing and scheduling delay.

However, when performing the first simulations, results show that for the single connectivity case, a very low load (200 kbps), and BLER target of 10%, the system was performing better in terms of delay than for the 1% BLER target simulations. After analysing PRB allocation, it was confirmed that 200 bytes payload size combined with the use of short TTI was causing packet segmentation for 1% BLER target case. With lower BLER target (1%) more robust MCS are used, which means that the coding rate will be lower. The increase of header bytes was sometimes forcing the scheduler to use two different TTIs to transmit a single packet, which was causing the delay to increase.

For that reason, payload size was changed to 50 bytes, which is also one of the possible sizes proposed in the simulation assumptions for URLLC in [1]. Figs. 26 and 27, show baseline results for a 50B payload size, where BLER target is set to 1% and the rest of the parameters are the same as in the previous sections. Results do not differ very much from results for 200B payload size simulations.

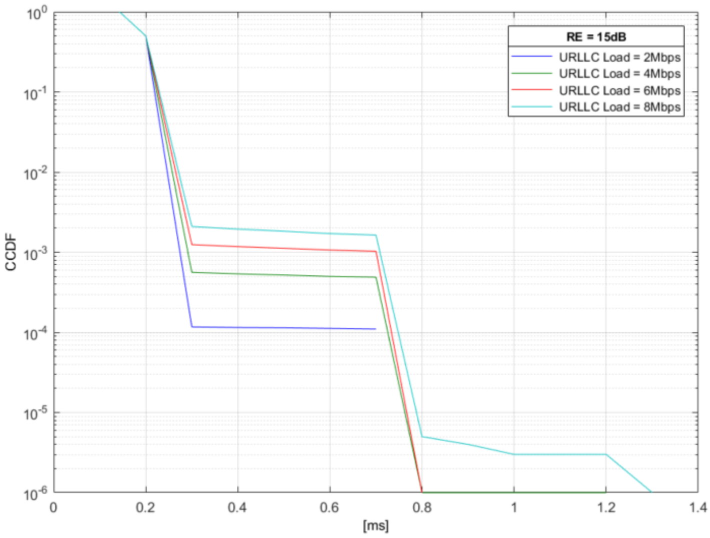

Figure 26. Payload Size = 50B - Delay Per FTP3 Packet for different loads with CRE = 15dB

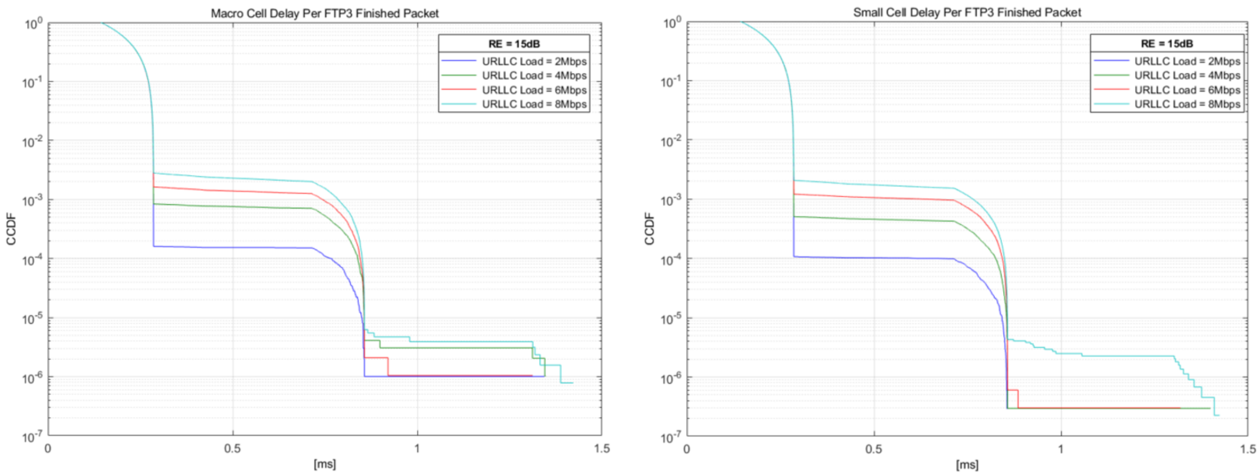

Figure 27. Payload Size = 50B - Per Layer Delay Per FTP3 Packet for different loads with CRE = 15dB

3.4. Multi-Connectivity Mode Configuration

3.4.1. Dual Connectivity in FREAC

This section explains how the dual connectivity algorithm is implemented in the system level simulator. A new feature for data duplication was added in the simulator for the development of this project. There is independent packet scheduling for each Component Carrier (CC). Duplication is performed at PDCP level. Fig. 28 shows a simplified modelling of the implementation. *FR_Traffic* class is roughly equivalent to PDCP layer, which is common for MgNB and SgNB. From it depends the *FR_DuplicationDC* class, which handles data duplication and dual connectivity. Then, as separate layers for MgNB and SgNB, there are FR_ConnL3, FR_ ConnL2, FR_RLC, etc., which correspond to the rest of the protocol stack layers such as RLC, MAC and PHY layers.

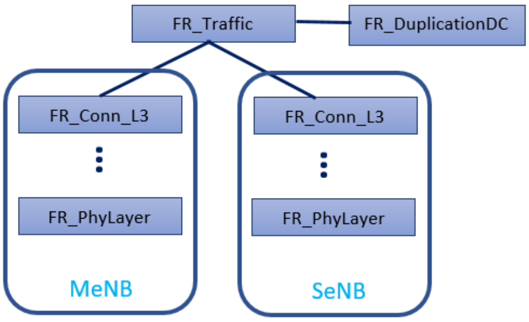

Figure 28. Simplified Model of DC in FREAC

Regarding the SgNB selection, it is split into two different steps. One is the Radio Resource Control (RRC), which is on the UE side, and the other is the SgNB configurations, which is on the base station side and is based on the measurements. This will be further explained in following sections.

3.4.2. Second Cell Selection Criteria

Aiming to optimize the number of UEs that should be configured with the dual connectivity mode, this research question tests the different conditions that could be imposed for a UE to configure it. The following-subsections explain the different studied options and the results obtained for each of the proposed solutions.

The idea is to configure dual connectivity only for those UEs that actually need it to meet the service requirements (mostly cell-edge UEs). It might exist the case where the UE is sufficiently close to the macro or the small cell that it can only be served by the primary cell (PCell). It has a strong serving RSRP and is in good channel conditions. That would be UEs 1 and 3 of Fig. 29. And it can also be the case where the UE does not have good enough channel conditions so it needs to configure a secondary cell (SCell) to improve its performance (UE2 in Fig. 29).

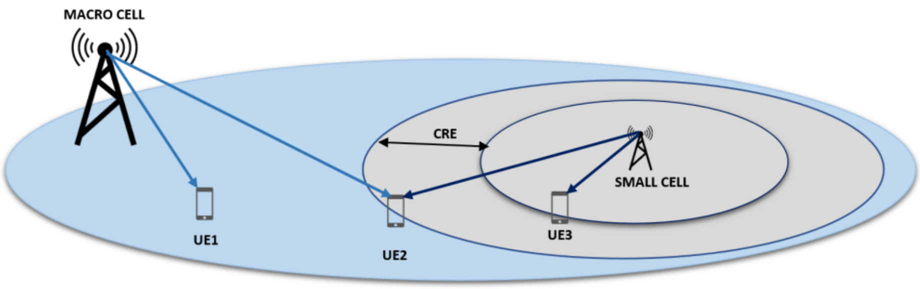

Figure 29. Connectivity options in the research scenario

In DC certain conditions are applied to decide if a secondary cell for the UE in question should or not be configured. These conditions are based on Measurement Reports triggering, and are part of what is referred to as LTE and NR Events in 3GPP. The criteria for triggering these events are evaluated after L3 filtering and must meet the requirements during a certain Time to Trigger (TTT) before the event is reported. For this study, an event based on LTE event A3 has been used. A UE needs to satisfy the following condition for the SCell to be configured.

Event A3 – Neighbour becomes offset better than PCell/PSCell

As explained in the previous subsection, the dual connectivity mode should only be configured for a UE that really needs it to meet the requirements for a certain service. Furthermore, it makes no sense to use resources of another cell if the requirements are already met with one. For that reason, the area of UEs that should enable DC must be controlled. UEs close enough to the macro cell to meet the requirements should never be in dual connectivity. Only UEs with small cell as serving cell should then enable DC. Parameters used below have to be redefined when comparing to Event A3 parameters defined in [19]. This is done to make the condition more understandable for the case.

$$RSRP_s < RSRP_m - CRE + DC_{Range} \ | \ m \in M \ and \ s \in S \qquad (3.2)$$

Where m or s is the corresponding cell a set of macro (M) or small (S) cells. CRE is the corresponding Cell Range Extension value, and DC_{Range} is the parameter that allows to control the size of what can be called *Dual Connectivity Area*, which would correspond to the green area in Fig. 30.

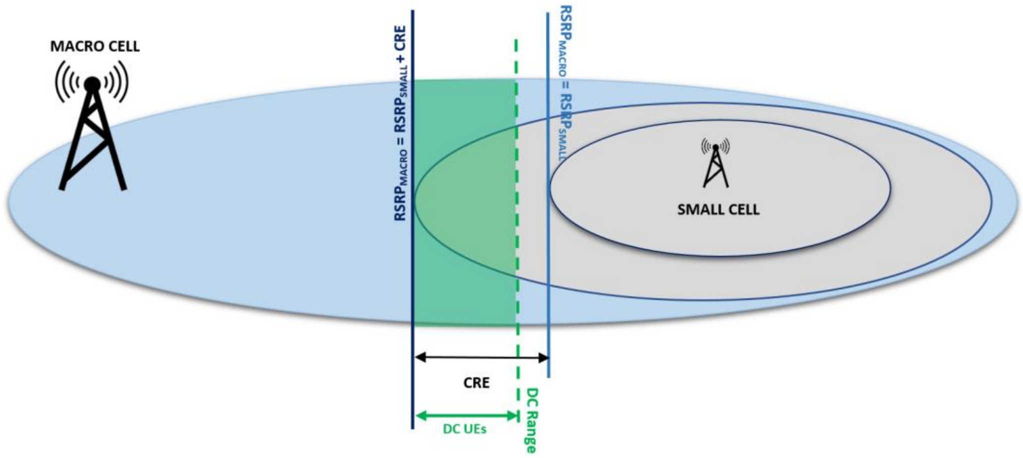

Figure 30. Dual-connected UEs Area

Contribution to FREAC

The option of enabling this condition was not included in the simulator. Therefore, the author needed to include it and properly test it before continuing with the following simulations.

3.5. Data Duplication Configuration

3.5.1. First Implementation

In the initial stage of the studies, data duplication implementation for dual connectivity performs blind data duplication. This means that each packet arriving to the MgNB is cloned and sent to the SgNB, and the latter, no matter what is happening in the PCell link, would send the packet. There is no coordination between the two gNBs, so each will send the packet to the UE as soon as possible. Scheduling is independent for both links, and so is the HARQ process.

As shown in Fig. 31 once the packet is received in the UE, an ACK or NACK is sent through the corresponding link depending on if the decoding of the packet was successful or not, respectively. The *Notify()* function is in charge of letting know upper layers if the packet has been successfully received. This function also controls that the sample added to the CDF function for the *'Delay Per FTP3 Finished Packet'* statistics is the corresponding to the first successful packet received at the UE. In that way, it is ensured that despite of duplicating in both links, only the sample for a certain packet that is first received is collected.

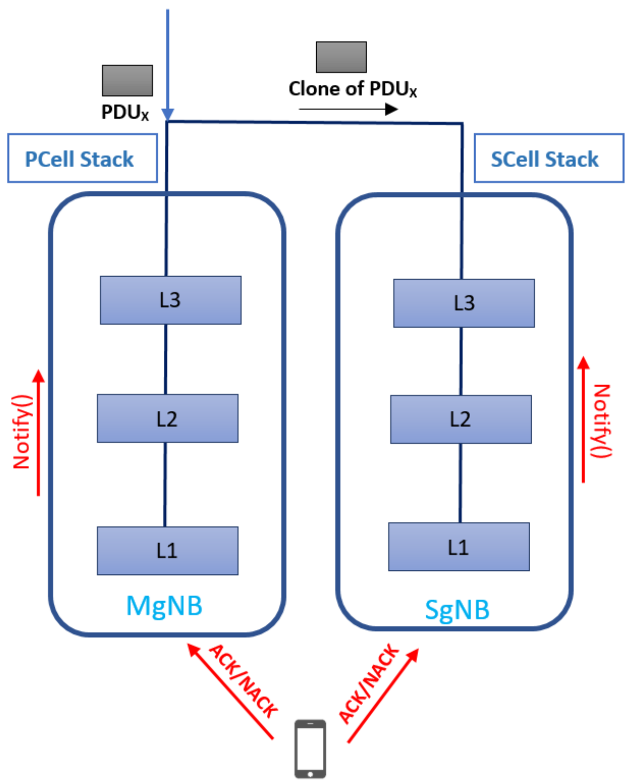

Figure 31. FREAC Data Duplication for DC Scheme

This is the simplest duplication that can be performed. It has several disadvantages. If, for example, a packet is sent through the first link and successfully decoded at the UE before the cloned copy of that packet is sent through the second link, that second transmission becomes unnecessary. The copy would be discarded at the UE, since that data is already available. And not only the second packet, but also the corresponding ACK, could be avoided, thus reducing unnecessary traffic creation in the network.

Duplicating every single packet that arrives to the MgNB and sending it through both links sometimes creates unnecessary traffic; and therefore, interference. This leads to a high resource utilization, so spectral efficiency is decreased. Hence, a feature that avoids unnecessary transmissions might improve network conditions.

3.5.2. **Optimization: PDU Cancelation**

To reduce the aforementioned unnecessary traffic and associated interference in the network, a new feature was tested in FREAC to avoid the unnecessary transmissions. This is referred to as *PDU Cancelation* feature. It consists on avoiding the transmission of a packet that has already been correctly received at the UE. As depicted in Fig. 32, if a packet sent through the MgNB is correctly decoded at the UE and an ACK is sent through the corresponding link, *Notify()* function (already mentioned in the previous subsection) value is shared with the SgNB. SgNB will use that information for the *CancelPDU()* function to cancel the PDU (i.e., not sending it), no matter at which layer the packet is. The PDU cancelation notification can happen in both directions. MgNB can cancel the PDU because SgNB has already successfully sent it, and vice versa.

Since the cancelation can be done at any layers (i.e. also below PDCP), it might be the case where part of the packet is already sent. For that case, the rest of the packet would be cancelled. Some statistics on packet discarding have also been extracted. Due to the fact that the PDUs are segmented in some layers, the only way of counting the discarded packets is per bytes. Section 4.3, shows the gain obtained by using this feature and includes some discarding statistics.

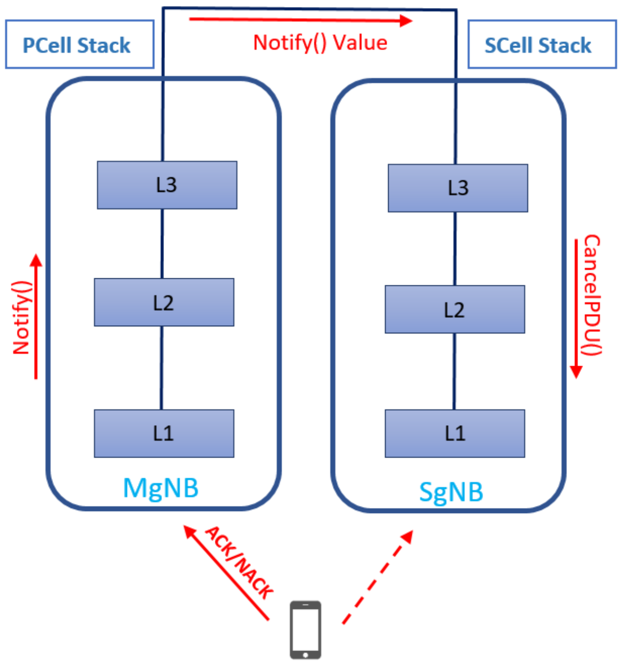

Figure 32. FREAC PDU Cancelation for Data Duplication DC

It was considered the option of acknowledging the packet received through one connection, also to the second connection. That is, sending the ACK for a certain packet identifier, to both links. However, since the links are not synchronized, and because of the PDU ID tracking currently implemented in the simulator, it was not possible to control what packet was being acknowledged in the second connection. For that reason, the algorithm previously explained was chosen.

4. Dual Connectivity Results

This section presents the results for the analysis of applying data duplication in DC for URLLC use case. Section 4.1 shows the results for a low load and low interference conditions scenario, were any gain is obtained from applying dual connectivity. eMBB traffic is added to study high interference conditions scenario in Section.4.2, and it is proved that gain is obtained from data duplication in DC. Finally, in Section 4.3 it is analysed the benefit of using PDU Cancellation for low and high URLLC offered loads.

4.1. Low Load and Low Interference Conditions

First simulations were performed at very low load (2Mbps). Fig. 33 shows the CCDF of URLLC radio latency for single and dual connectivity (DC), for both, 1% and 10% BLER targets for first transmissions. For DC mode, two different values where simulated for the *DCRange* to vary the rate of UEs using DC: 10 and 15dB (covering with the latter the case where the *Dual Connectivity Area* covers the whole CRE area).

Two main observations can be extracted from the figure. The first conclusion that can be noticed is that the performance obtained for single connectivity with 1% BLER target is very similar to the performance at 10% BLER target. The second observation is that single connectivity performs noticeably better than dual connectivity as the *DCRange* increases. Both observations are due to the same reason: the limited inter-cell interference level in this scenario due to the low load conditions. Although lack of interference has typically a positive impact on the system performance, it is not a beneficial condition to proof the benefit of data duplication for DC.

In the single connectivity case, for both, 1% and 10% BLER targets, the interference level is so low, that the system performs well below its BLER target, which makes it difficult to see any gain from reducing it further. Moreover, SC is performing so good, that enabling DC for the same traffic conditions just worsens the performance. As mentioned earlier, DC should only be enabled when needed. When enabling DC for some users, the traffic in the network increases due to packet duplication, which increases the interference and therefore, reduces the UE SINR which in turn results in more transmission errors and a larger retransmission rate, and thus increased delay. For this reason, increasing the range of DC use in the cell, and with that, increasing the number of dual-connected UEs, affects negatively the URLLC delay performance.

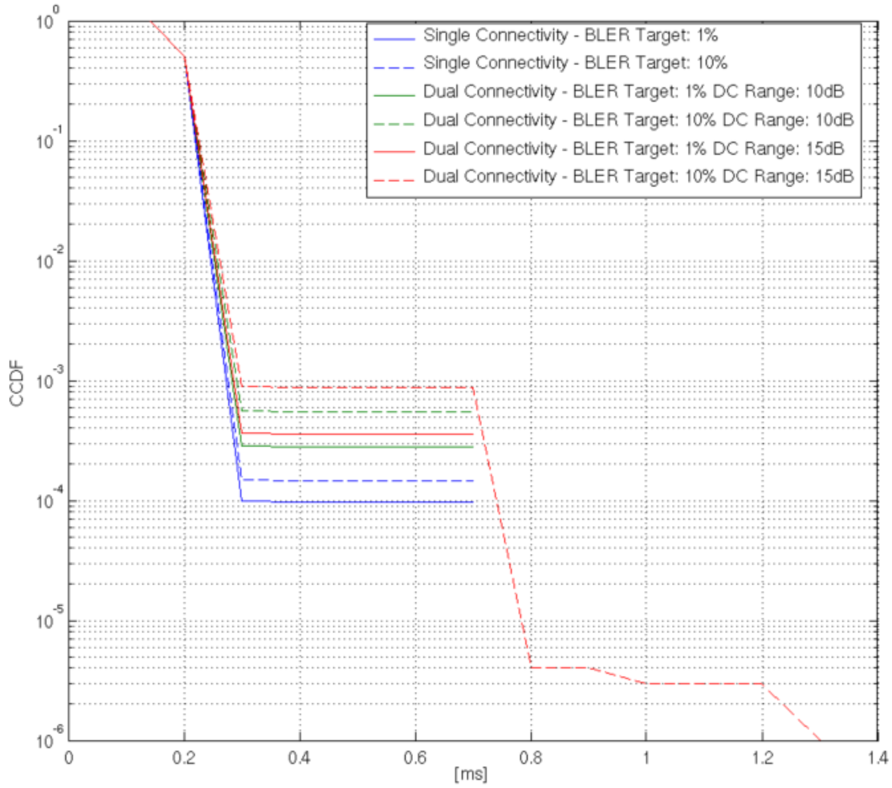

Figure 33. - Low Load and Interference Conditions - Delay Per Finished FTP3 Packet – BLER Target: 1%

As a conclusion, for low URLLC offered load and low interference conditions there is no gain on applying dual connectivity in the studied scenario. For this reason, and to study a more realistic scenario, we need to consider a scenario with higher traffic load, especially from eMBB users.

4.2. High Interference Conditions

To increase the interference level in the scenario, full buffer background eMBB traffic was added to the simulations. New UEs generating eMBB traffic only were added to the scenario.

Considering that all the base stations have now full buffer background eMBB traffic, and are operating at close to 100% load in terms of PRB usage, fulfilling URLLC requirements becomes more challenging. For this reason, lower URLLC loads have been considered for these simulations. Therefore, simulations were performed for loads between 500kbps and 6Mbps. At any load level, three different *DCRange* values were simulated: 5 dB, 10dB and 15dB, respectively. This allows seeing the effects of covering the different parts of the *Dual Connectivity Area*. Based on the learning from the results in Section 4.1, the BLER target is set to 1% and URLLC packets have priority over eMBB packets as they are latency critical.

Fig. 34 shows the CCDF of the radio latency of URLLC under the different offered URLLC loads for the single connectivity case. Performance is considerably similar in terms of delay for URLLC offered loads up to 4Mbps. Due to the full buffer background eMBB traffic, the UE SINR remains the same irrespective of the load. For this reason, and since the CRE is optimized for a load balance between the macro and small cells, there is not a big difference in performance until, at some point when increasing URLLC load further, the resource availability for instant scheduling of URLLC becomes limited. This results in queueing delays, and creates a bottleneck in the system where the URLLC latency increases significantly. The difference in this scenario can be appreciated for URLLC loads from 6Mbps. In Annex C, the effect of reducing the CRE can also be seen, which would increase the UEs on the macro cell (which cannot benefit from DC) and create a bottleneck at the macro layer. In that situation it is also possible to see the difference in performance between the different URLLC loads.

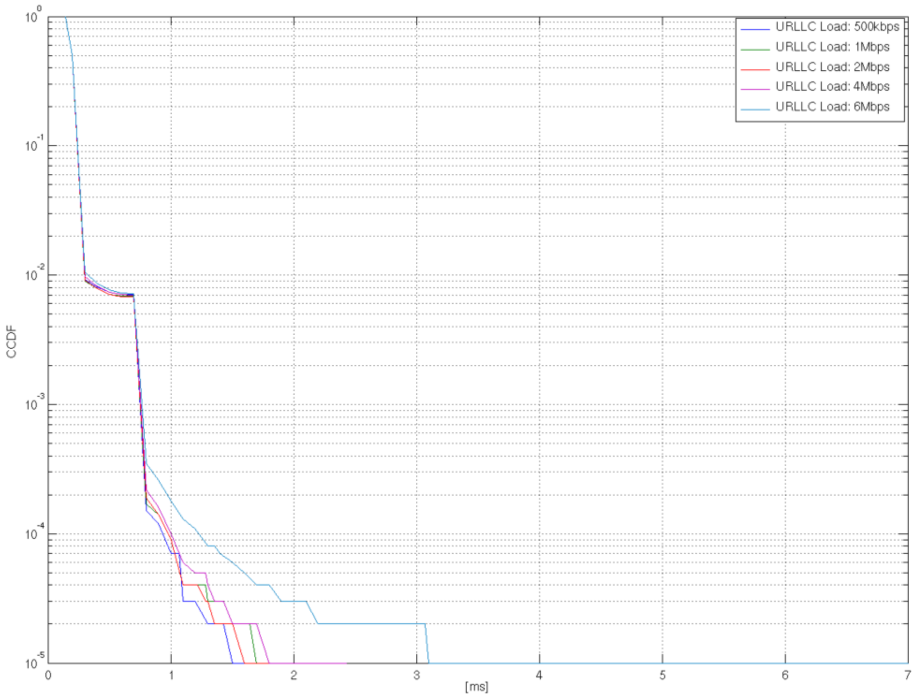

Figure 34. SC Delay Performance for different URLLC loads with eMBB traffic

Simulations were performed for both 1% and 10% BLER targets. Figures comparing results for the two BLER targets can be seen in Annex C. There can be seen that there is a mismatch when comparing to theoretical analysis shown in Section 2.4.1.2. It was expected to obtain good performance with a 10% BLER target in dual connectivity since there is a backup link. The plots extracted from the system-level simulator data present the average BLER. In reality, not all the transmissions will have the same BLER, however this assumption has not been considered in the theoretical analysis.

Even with 1% BLER target it is hard to meet the URLLC requirements, 10% BLER target results are not shown in the following plots.

Table 5 shows the approximated (it might vary depending on how the UEs are dropped in the scenario) percentage of the number of dual-connected UEs for the *DCRange* values simulated. These values are extracted from the 6Mbps URLLC load case, however they will remain the same irrespective of the load. This is because the RSRP does not vary with the load, and the DC setup conditions are based on such measurement. As expected, when *DCRange* value is increased, since the *Dual Connectivity Area* is also increasing, so does the number of UEs using DC.

Table 5. Number of DC UEs per DCRange setting

DCRange (dB)	~ #Dual-Connected UEs (%)
5	9%
10	19%
15	30%

Fig. 35 shows results obtained for 500kbps URLLC offered load and the different *DCRange* values. For this case, the best performance is obtained with *DCRange* = 15dB (covering the whole CRE area). When the number of DC UEs increases, the number of retransmissions decrease. This is the effect of having two links transmitting the same packet. If the packet is not correctly received through one of the links, it might be successfully sent through the other. Under this low URLLC load level, there is not queueing delay for URLLC packets since they have priority against the eMBB packets. Resource availability for URLLC traffic is high, so it is beneficial to allow a larger number of UEs to DC.

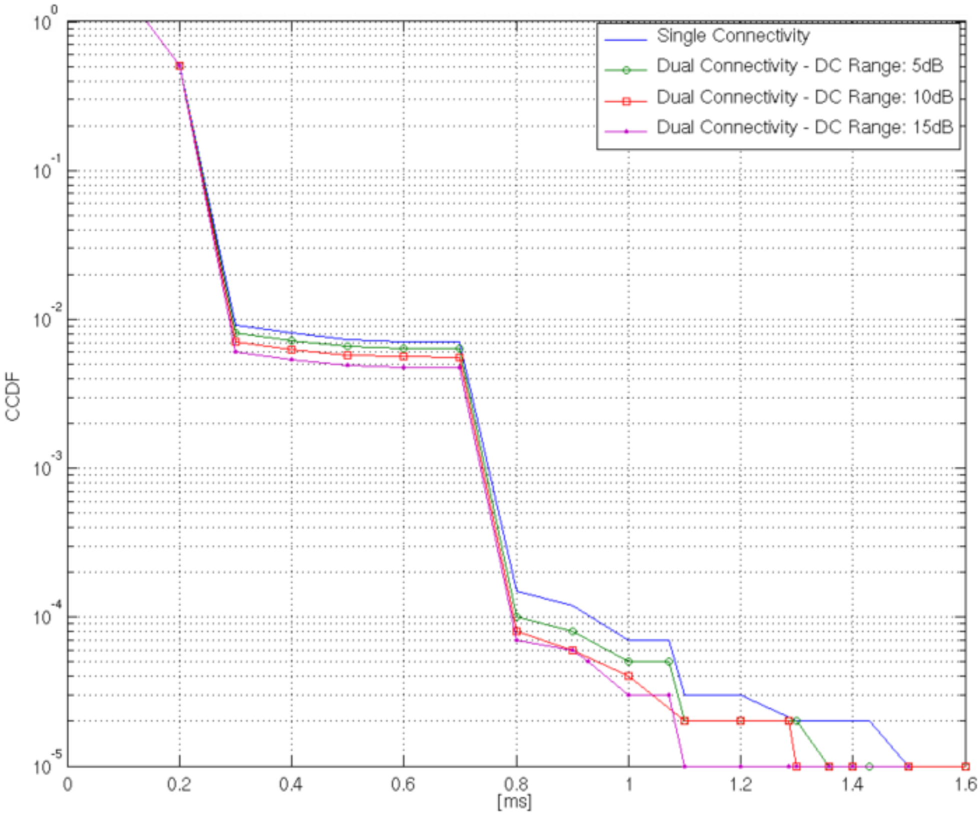

Figure 35. SC vs. DC Delay Performance for 500kbps URLLC load combined with eMBB traffic for different DC Range values.

However, as URLLC load increases, DC is not so beneficial, and the optimal value of *DCRange* is lower. Figs. 36 and 37 show results for 1Mbps and 2Mbps URLLC offered loads, respectively. The number of retransmissions keeps decreasing as the *DCRange* value increases. Nevertheless, for these cases, it starts to be appreciated how queueing delay slightly increases due to a higher resource occupancy. Since the number of packets arriving to the gNBs is higher, resource availability to schedule all that URLLC packets is lower. This leads to an increase of queueing delay. Fig. 38 shows version of the 2Mbps with the previous plots scaled to allow a better comparison. The case without modifications on the scale will be also shown from now on to allow the reader to observe the gain that dual connectivity offers. It can be clearly appreciated how the packets with the longest delays are removed from the CCDF with dual connectivity.

The interest of this study is to reduce as much as possible the chances of a packet exceeding the 1ms latency target for URLLC. For this reason, the optimal *DCRange* value is the one which ensures the lowest probability of that happening at 1ms latency. For 1Mbps and 2Mbps URLLC offered loads, the optimal *DCRange* decreases to 10dB. Although there is higher resource occupancy, there are still available resources for URLLC traffic that allow pushing more UEs to DC and obtaining benefit from it. Decreasing too much *DCRange* does not take any profit from these available resources. For this reason, *DCRange* = 5dB case is performing worse than *DCRange* = 10dB for 1Mbps and 2Mbps.

It can be seen that gain is very sensitive to the configuration and the load conditions. The reader is encouraged to refer to Annex B to obtain a per user analysis of the 2Mbps URLLC load with additional eMBB traffic case. There can be observed how performance improves for some of the UEs but worsens for some others.

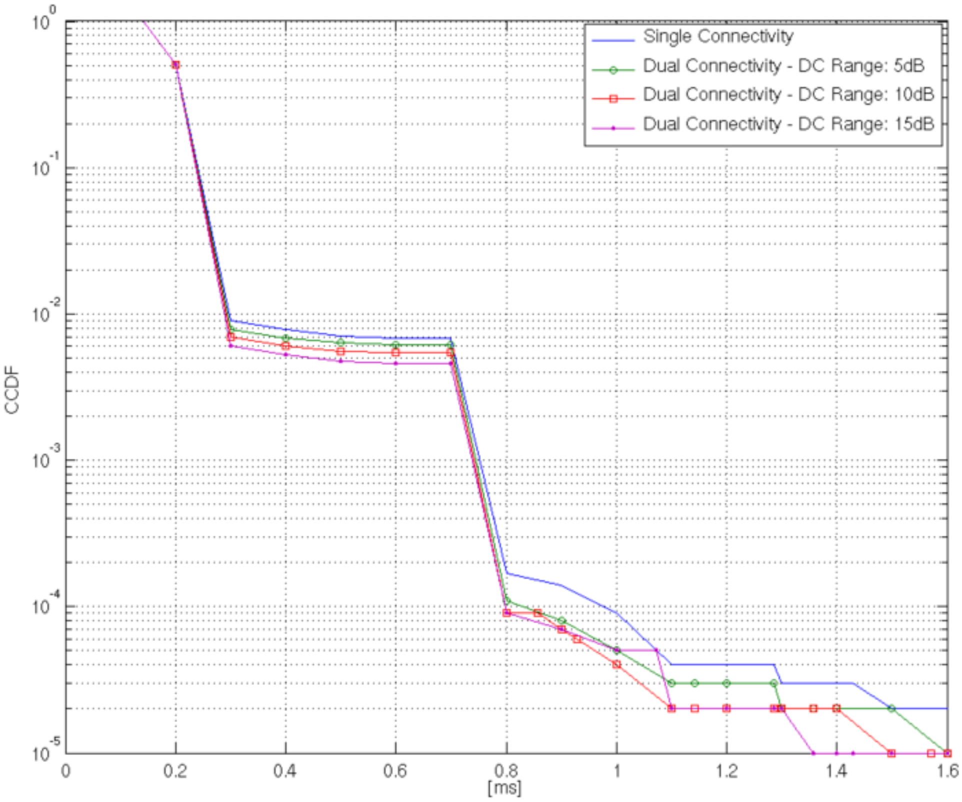

Figure 36.SC vs. DC Delay Performance for 1Mbps URLLC load combined with eMBB traffic for different DC Range values.

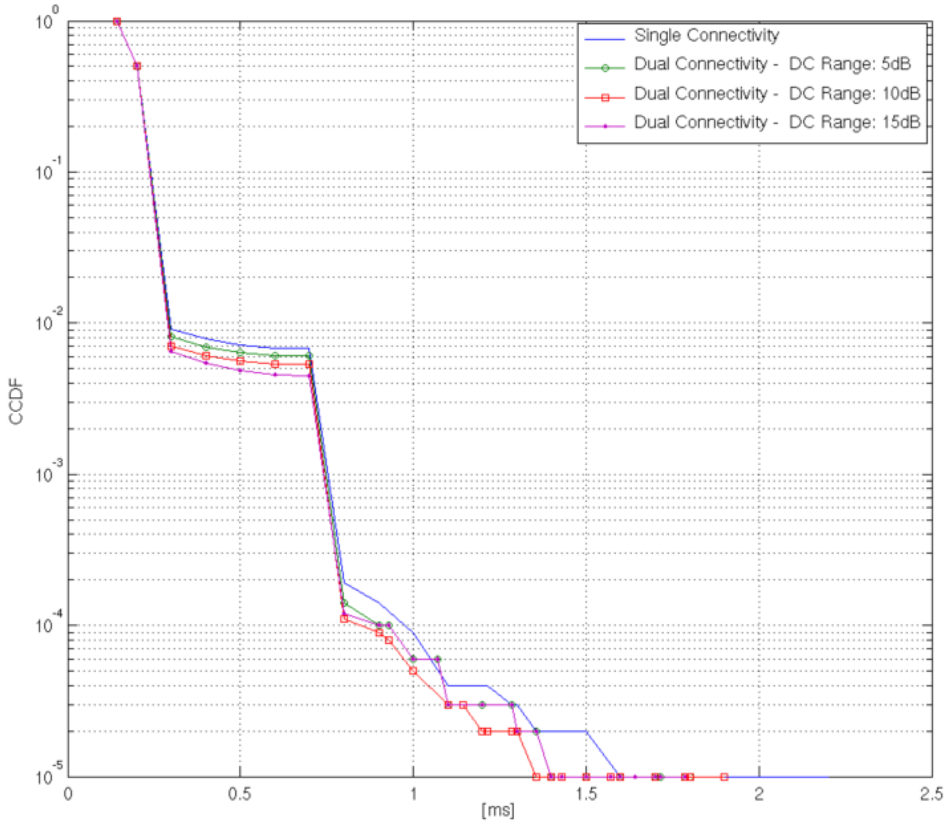

Figure 37. SC vs. DC Delay Performance for 2Mbps URLLC load combined with eMBB traffic for different DC Range values.

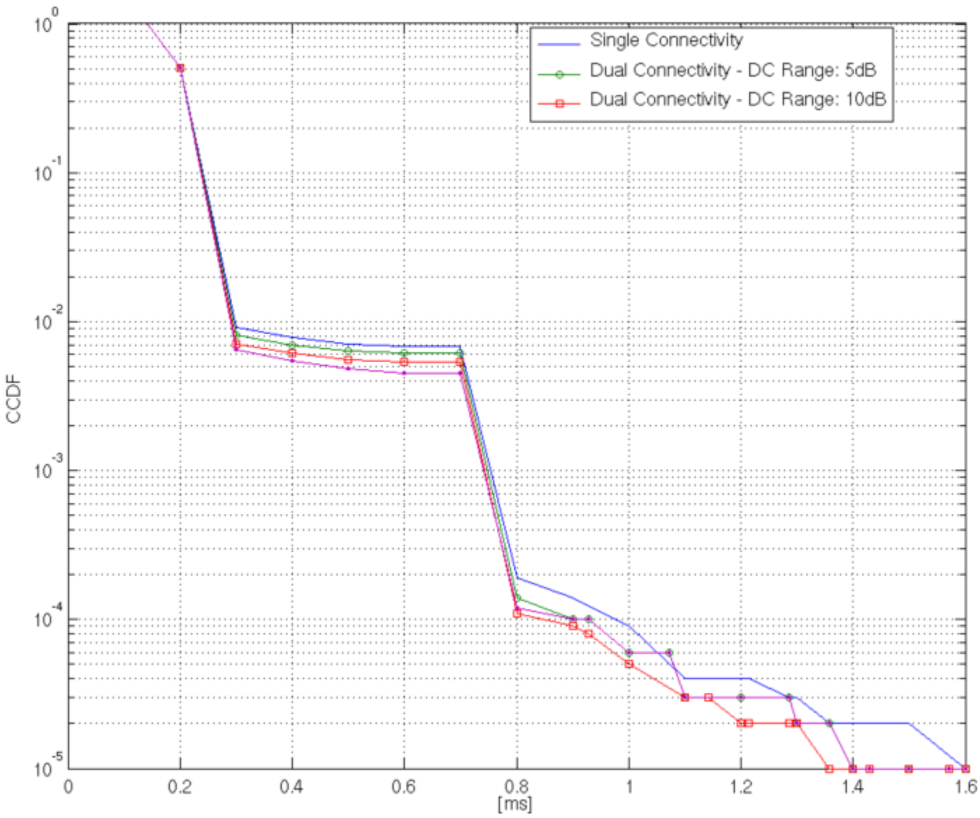

Figure 38. Zoom of Fig. 37.

Although URLLC requirements (i.e. 1 ms at 10^{-5} outage) are not met even at the lowest URLLC offered load (500kbps), simulations for higher loads were also performed to observe the effect of dual connectivity in cases with higher queueing delay. Figs. 39 and 40 show results for 6Mbps URLLC offered load. When increasing the URLLC load, *DCRange* optimal value keeps decreasing. It can already be observed in the figures how the lines for DC are completely crossing and there is higher gain on applying dual connectivity when *DCRange* = 5dB. It can also be observed in these two figures how DCRange = 15dB performs even worse than SC. This is caused by the excess of traffic in the network. Too many UEs in DC are creating such a high queueing delay that DC is not beneficial anymore.

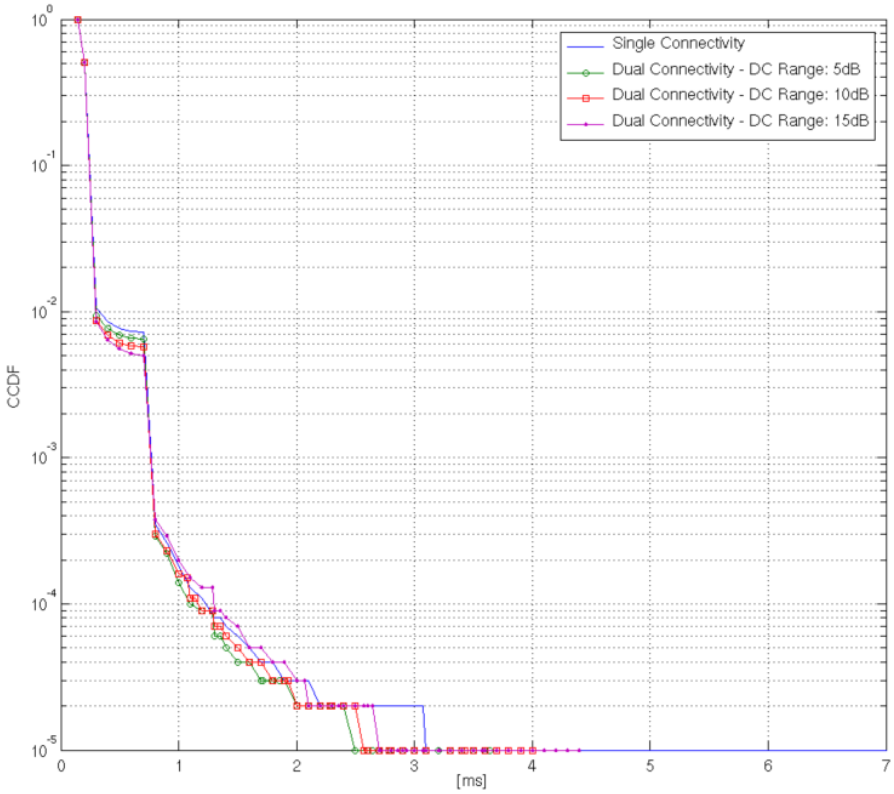

Figure 39. DC Delay Performance for 6Mbps URLLC load combined with eMBB traffic for different DC Range values.

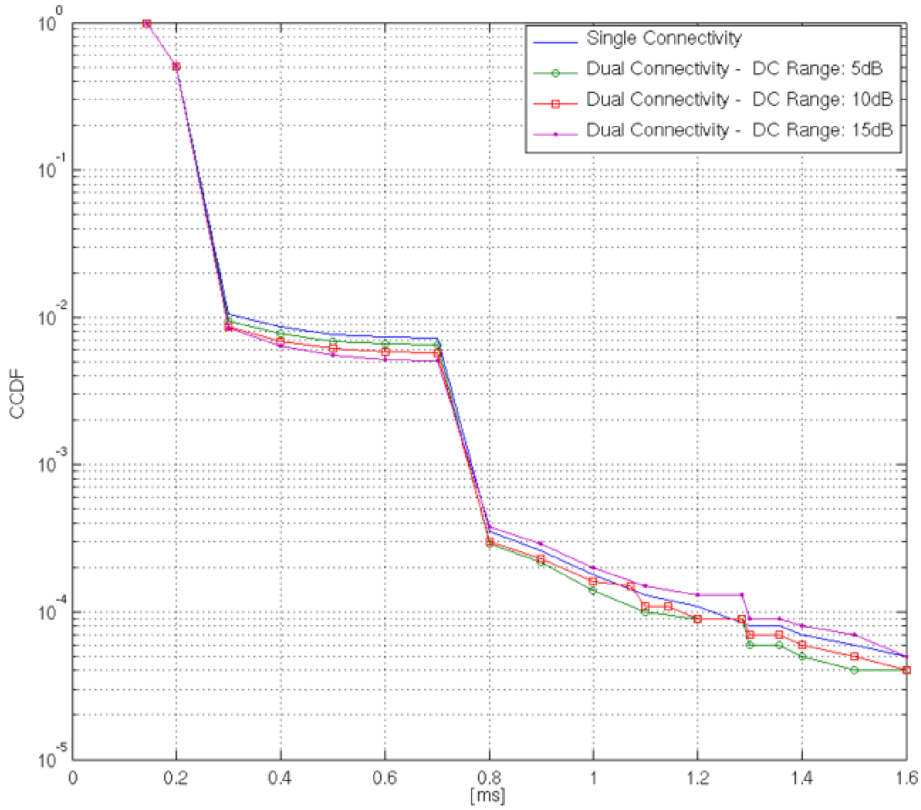

Figure 40. Zoom of Fig. 39.

It can be seen by any figure presented in this section, that dual connectivity provides gain specially on the tails of the latency CCDFs as expected. This is because the packets that were creating those tails in single connectivity might have been sent successfully through the *backup link*.

URLLC requirements are far away from being met for these cases, but it is proved that DC reduces the latency, especially in a high load and high interference scenario. There are scenarios that require very low latency and high reliability but with more relaxed requirements than the targeted in this project (examples can be found in [21]). For those cases, DC might be a potential solution to reduce latency.

Contribution to FREAC

eMBB traffic was already available in the simulator, but for implementation reasons, it was not compatible with enabling dual connectivity. Some modifications in the code were needed to allow the use of eMBB full buffer background traffic when enabling DC. Also new functions and conditions were added to the code to avoid eMBB UEs to enable DC mode, since the effect of using DC wanted to be tested for URLLC UEs only.

4.3. Dual Connectivity Optimization: PDU Cancelation

As explained in Section 3.5.2., PDU cancelation at the network side was implemented as an optimization of the data duplication for dual connectivity algorithm. Since this optimization is supposed to avoid unnecessary transmissions through one of the links once the packet has been successfully received through the another, some gain in delay performance is expected from applying this optimization.

Most of the packets that PDU Cancelation feature avoids sending are those that are in the buffer of the node. This means that this PDU Cancelation feature is expected to be especially beneficial for scenarios with high loads, where the queueing delay is also high. For low loads, discarding rate should be so low, that there should not be any gain effect obtained from this optimization. On the other hand, for high loads, discarding rate should increase.

Performance is very similar for some of the different URLLC loads simulated. Therefore, results shown in this section are only shown for 500kbps, 6Mbps and 8Mbps.

Fig. 41 shows the delay performance for 500kbps URLLC load, dual connectivity with PDU Cancelation enabled. For this case, the *DCRange* trend remains the same as when PDU Cancelation feature is not enabled. Fig. 42 shows the difference for the same load, between SC, DC and DC with PDU Cancelation. The gain from applying cancelation is very small due to the low percentage of cancelation.

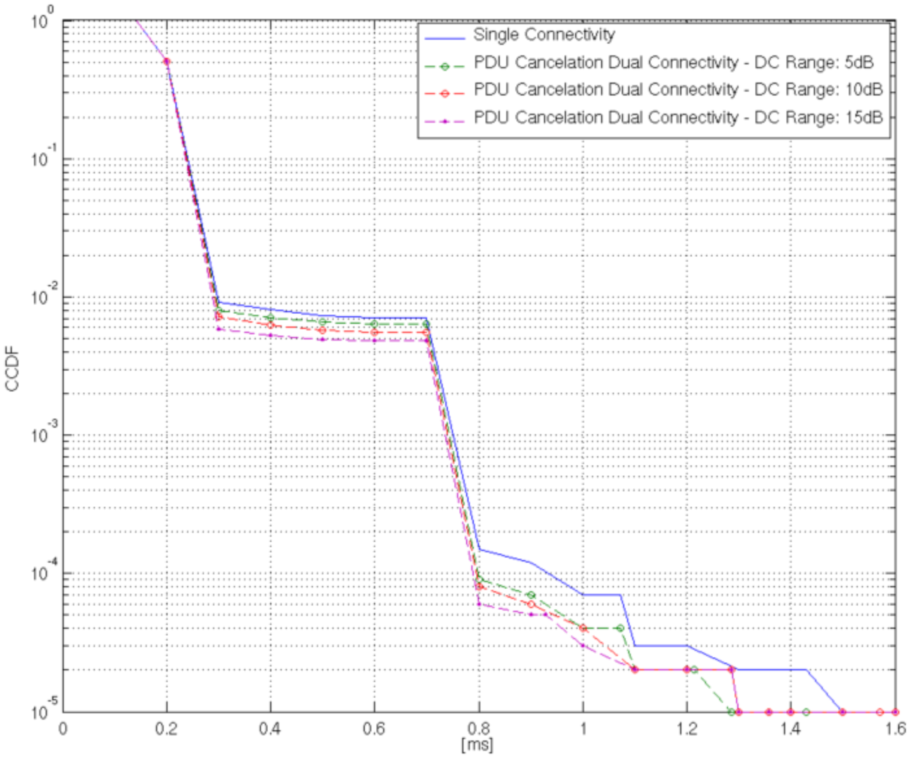

Figure 41. PDU Cancelation Delay Performance for 500kbps URLLC load combined with eMBB traffic for different DC Range values.

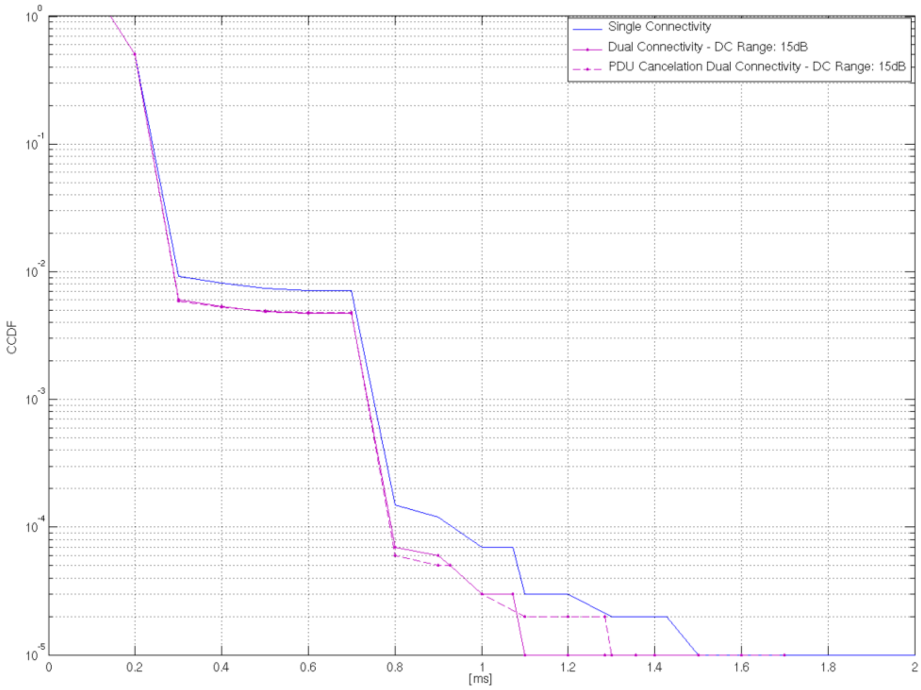

Figure 42. SC vs. DC with PDU Cancelation Delay Performance for 500kbps URLLC load combined with eMBB traffic for optimal DC Range value.

As observed in Fig. 43, for this case, optimal *DCRange* value remains also the same as for DC without PDU Cancelation. However, it can be appreciated how lines for the different DCRanges start to be close. As previously mentioned, increasing the load up to 6Mbps creates a considerable queueing delay that allows appreciating the gain on applying PDU Cancelation for DC. This can be observed in Figs. 44 and 45.

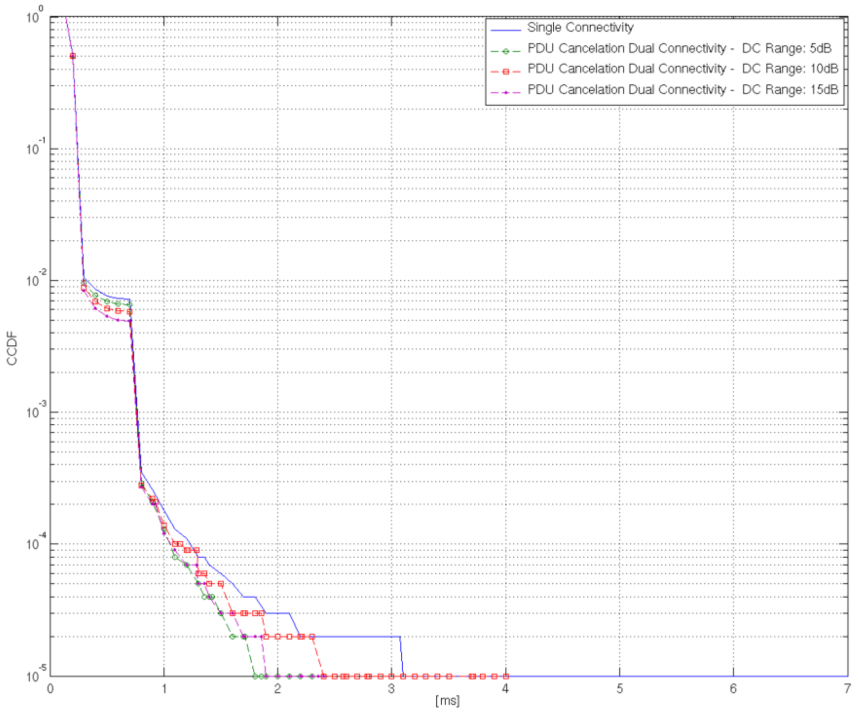

Figure 43. PDU Cancelation Delay Performance for 6Mbps URLLC load combined with eMBB traffic for different DC Range values.

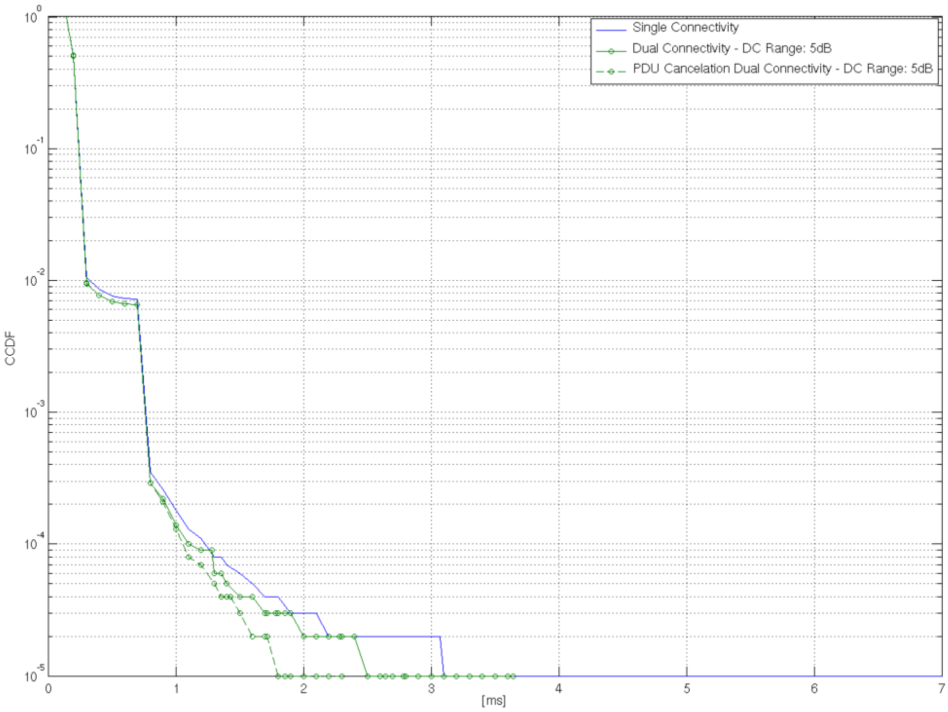

Figure 44. SC, DC and DC with PDU Cancelation comparison for 6Mbps URLLC Load combined with eMBB traffic.

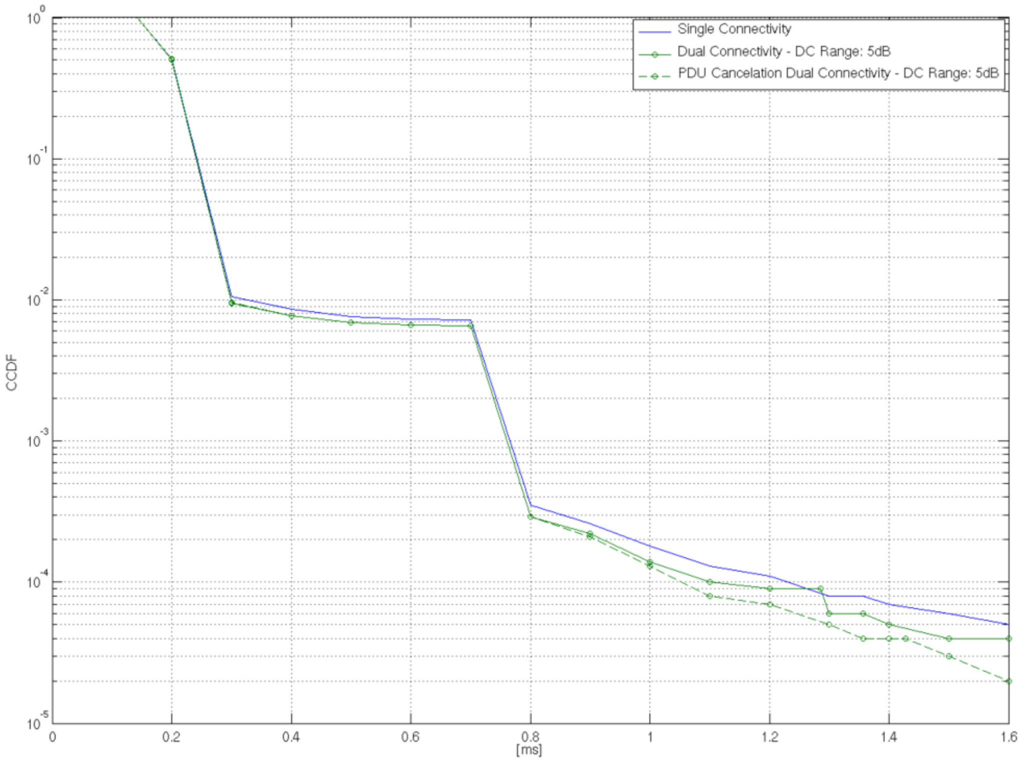

Figure 45. Zoom of Fig. 44.

As the load keeps increasing, the optimal *DCRange* value could change with respect to the DC case without PDU Cancelation. The higher the *DCRange* value, the higher the number of dual-connected UEs is. The higher the queueing delay, it takes more benefit from PDU Cancelation feature. The gain might be such, that it allows pushing more UEs to DC. .

5. Conclusions and future development

The main objective of this study was to increase supported URLLC offered load by means of DC with PDCP-level duplication and understand the factors that affect its performance. The study is done through system-level simulations, with high level of realism in the NR radio access technology assumptions.

Since the packet is sent twice through two different links some gain is expected in the radio latency. However, data duplication has a cost in terms of resource utilization. Higher resource utilization leads to a higher interference level, lower SINR and therefore, lower MCS, which may lead to increased queueing delay if data duplication for DC is not operated carefully.

To deal with this issue, two different means are studied in this investigation. First, controlling the number of UEs using DC. Increasing or reducing the area where the UEs will enable DC depending on the load helps to optimize resource utilization. Furthermore, to avoid wasting resources, the network is enforced to discard packets that are already received at the UE.

The studied scenario is a heterogeneous network of 21 macro cells with 4 pico cells cluster per macro cell, and 3D channel model. The first conclusion that can be drawn is that the studied HetNet scenario, with small cells clustered under the macro layer, is able to support up to 8Mbps URLLC offered load with SC mode when no background eMBB traffic is present. These are new and improved results compared to previous SC studies presented e.g. in [16], which supported, for a 21 macro cell scenario and a 2D channel model, up to 2Mbps.

Under low URLLC load and low interference conditions, SC is therefore sufficient to fulfil URLLC requirements in this scenario. Enabling DC with data duplication only worsens the interference conditions because of the additional traffic created with the duplicated packet copies and so, the URLLC delay performance is worse than with SC.

Adding full buffer background eMBB traffic to the previous scenario not only shows the performance under some more realistic conditions but also determines an improvement in the latency that can be achieved with DC. However, this gain is very sensitive to DC mode configuration and scenario conditions. Furthermore, even for the low load case (i.e., 500kbps) URLLC target is not met.

The obtained gain is largely sensitive to the conditions of the UE and the network when applying DC. Moreover, most of the UEs with delay longer than 1ms have a small cell as serving cell. This is likely be due to the physical proximity between the small cells, which operate in the same frequency layer, and hence, create a strong interference coupling among them. Some type of interference cancelation mechanism could be needed to alleviate the issue.

We have also investigated PDU Cancelation, i.e. the network discard of copies already received at the UE, and showed that it provides additional gain respect to blind duplication in DC for high loads only, where the duplication rate is high and queueing delay can be avoided by the cancellation.

There are several 5G applications with more relaxed latency demands compared to the 1ms latency requirement with 99.999% reliability requirement of URLLC. Therefore, DC could be used to improve the delay performance for other 5G applications with a more relaxed time restriction.

In the light of findings, DC with data duplication still considered as a promising option to meet the URLLC requirements (1ms latency requirement with 99.999% reliability). As future work, there are many options yet to be investigated to improve the solution further. First, Inter-Cell Interference Coordination (ICIC) technique could be used to reduce inter-interference between small cells. Also, the case when multi-connectivity is used and duplicated packets are sent through more than two nodes should be investigated.

Specially focusing on configuration for data duplication, there are several optimization options. The current design is such that, duplication is performed through all the configured links (two, for the case of DC). Although there is cancelation of a copy in case the packet is correctly received through one of the links, the packets are initially duplicated and intended to be sent through both links without any kind of intelligence. Two possible optimizations are proposed as future developments:

- Always duplicate the packet arriving through the MgNB but use only the best serving cell within all the configured nodes. Based on load of the base stations and channel conditions, send the packet only through the best link.
- Dynamically determine the need or benefit of duplication. Depending on channel conditions, decide if for a certain time, if the UE is going to benefit or not of data duplication.

6. Bibliography

[1] 3GPP, ""Technical specification group radio access network; Study on new radio (NR) access technology physical layer aspects," TR 38.802, Rel-14 V.1.1.0, 3GPP, Feb. 2017.

[2] 3GPP, "Technical specification group radio access network; Study on scenarios and requirements for next generation access technologies," TR 38.913,Rel-14 v.14.1.0, 3GPP, Dec. 2016.

[3] B. Soret et al., "Fundamental Trade-offs among Reliability, Latency and Throughput in Cellular Networks," IEEE Proc. GLOBECOM Wkps., Dec. 2014, pp. 1391–96

[4] Ericsson, "On URLLC design principles," 3GPP, TSG-RAN WG1 #86bis R1-1609634, Lisbon, Portugal, Oct., 2016.

[5] 3GPP, "Technical specification group services and system aspects; Feasibility study on new services and markets technology enablers for critical communications; Stage 1," TR 22.862, Rel-14 V.14.1.0, Sept. 2016.

[6] D. Ohmann, A. Awada, I. Viering, M. Simsek and G. P. Fettweis, "Impact of Mobility on the Reliability Performance of 5G Multi-Connectivity Architectures," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.

[7] D. López-Pérez, D. Laselva, E. Wallmeier, P.Purovesi, P. Lundén, E. Virtej, P. Lechowicz, E. Malkamaki, M. Ding. "Long Term Evolution-Wireless Local Area Network Aggregation Flow Control" *IEEE Access, vol. 4,,* pp. 9860-9869, 2016.

[8] 3GPP T. 36.872, "Technical Specification Group Radio Access Network; Small Cell enhancements for E-UTRA and E-UTRAN - Physical layer aspects (Release 12)," December 2013.

[9] Wang, H., Rosa, C., & Pedersen, K. I. (2016). Dual connectivity for LTE-advanced heterogeneous networks. Wireless Networks, 22(4), 1315–1328. DOI: 10.1007/s11276-015-1037-6.

[10] 3GPP, "LTE, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical layer, Measurements" TS 36.214 V.13.0.0. Rel.13. 2016.

[11] IMT Vision – "Framework and overall objectives of the future development of IMT for 2020 and beyond", International Telecommunication Union (ITU), Document, Radiocommunication Study Groups, February 2015.

[12] E. Dahlman, et.al., "5G Wireless Access: Requirements and Realization", IEEE Communications Magazine – Communications Standards Supplement, December 2014.

[13] C. Rosa et al., ‘‘Dual connectivity for LTE small cell evolution: Functionality and performance aspects,’’ IEEE Commun. Mag., vol. 51, no. 6, pp. 137–143, Jun. 2016.

[14] Wireless World Reserch Forum. Communication Architectures and Technologies – Working Group C. ‘‘LTE Small Cell Enhancement by Dual Connectivity”. Nov. 2014.

[15] G. Pocovi, B. Soret, K. I. Pedersen, P. Mogensen. “MAC Layer Enhancements for Ultra-Reliable Low-Latency Communications in Cellular Networks”. IEEE International Conference on Communications Workshops (ICC), 2017.

[16] C.Li, J. Jiang, W. Chen, T. Ji, J. Smee. “5G Ultra-Reliable and Low-Latency Systems Design”. Qualcomm Research, Qualcomm Technologies, Inc.

[17] G. Pocovi, K. I. Pedersen, P. Mogensen, B. Soret. “Radio Resource Management for 5G Ultra-Reliable Low-Latency Communications” from Doctoral dissertation: G.Pocovi. “Radio Resource Management for Ultra-Reliable Low-Latency Communications in 5G

[18] G. Pocovi, K. I. Pedersen, B. Soret, Mads Lauridsen and P. Mogensen, “On the impact of multi-user traffic dynamics on low latency communications", International Symposium on Wireless Communication Systems (ISWCS), Sept. 2016.

[19] 3GPP, “LTE, Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC), Protocol Specification” TS 36.331 V.13.7.1. Rel.13. 2017.

[20] A. Aijaz. “Packet Duplication in Dual Connectivity Enabled 5G Wireless Networks: Overview and Challenges”. Apr. 2018.

[21] 3GPP,”Technical Specification Group Services and System Aspects; Service requirements for the 5G system”. TS 22.261 V16.3.0. Rel. 16. 2018.

Glossary

BLER	Block Error Rate
BS	Base Station
CA	Carrier Aggregation
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CN	Core Network
CRE	Cell Range Extension
CT	Core Network and Terminals
CQI	Channel Quality Indicator
DL	Downlink
DC	Dual Connectivity
eMMB	Enhanced Mobile Broadband
gNB	GNode-B
HARQ	Hybrid Automatic Repeat Request
ICIC	Inter-Cell Interference Coordination
LTE	Long Term Evolution
MAC	Medium Access Control
MC	Multi-Connectivity
MCS	Modulation Coding Scheme
MgNB	Master gNode-B
mMTC	Massive Machine Type Communication
NR	New Radio
OFDM	Orthogonal Frequency Division Multiplexing
PDCP	Packet Data Convergence Protocol

PDU	Protocol Data Unit
PRB	Physical Resource Block
RAN	Radio Access Network
RAT	Radio Access Technology
RRC	Radio Resource Control
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Reference Signal Strength Indicator
RTT	Round Trip Time
SA	Service and Systems Aspects
SC	Single Connectivity
SDAP	Service Data Adaptation Protocol
SgNB	Secondary gNode-B
TCP	Transmission Control Protocol
TTI	Time To Interval
TTT	Time To Trigger
UE	User Equipment
UL	Uplink
URLLC	Ultra-Reliable Low Latency Communications

Annex A

This annex contains more specific information related to parameters configuration for the studied scenario.

Table 6. Specific Scenario Parameters

Parameters	Macro layer-NR	Pico layer -- NR
Layout	7 sites, 21 cells, wrap around	4 picos per cell
Inter-BS distance	500m	cluster
Carrier frequency	2 GHz	3.5 GHz
Simulation bandwidth	10 MHz	
BS power	46 dBm	30 dBm
Pathloss Model	3D-UMa PL = $22.0\log_{10}(d_{3D}) + 28.0 + 20\log_{10}(f_c)$ LOS Probability	3D-UMi PL = $22.0\log_{10}(d_{3D}) + 28.0 + 20\log_{10}(f_c)$ LOS Probability
Shadowing	6dB	
Antenna Height	32m	10m
UE Antenna Height	1.5m	
Antenna gain	18dBi	5dBi
UE Antenna gain	0dBi	
Antenna configuration	2x2 cross-polar	
UE dropping	2/3 UEs randomly and uniformly dropped within the clusters, 1/3 UEs randomly and uniformly dropped throughout the macro	
Radius for small cells dropping in a cluster	50m	
Radius for UE dropping in a cluster	70m	
Minimum distance	Small Cell – Small Cell: 20m	

(2D Distance)	Small Cell – UE: 5m Macro – Small Cell Cluster Centre: 105m Macro – UE: 35m Cluster Centre – Cluster Centre: 100m	
BS antenna pattern	TR36_814	Isotropic
BS antenna height	32 m	10 m
Subcarriers/RB BS	12	
Subcarrier spacinSswg(kHz)	15	
Symbol/Second	14000	
Cell Selection Criteria	RSRP w/ 20dB CRE	
SCell selection event	N/A	A4-RSRQ (-15dB)
Receiver Type	LMMSE_IRC	
Traffic Model	FTP3: based on FTP model 2 with the exception that packets for the same UE arrive according to a Poisson process and the transmission time of a packet is counted from the time instance it arrives in the queue.	
LLC Transport Type	UDP	
Maximum PDU Size	1500B	
Link Adaptation	Outer Loop Link Adaptation (OLLA) algorithm 1% initial BLER target for LLC	

Annex B

This annex contains a per user analysis for the case for a simulation with 2Mbps URLLC offered load per macro cell combined with full buffer eMBB traffic. CRE is set to 15dB, as well as DC Range, which is also set to 15dB. There are 840 UEs, 66% of which are URLLC UEs, while the remaining 33% are MBB UEs.
Table 7 presents parameters for all the URLLC users whose delay during the simulation is above 1ms. That is, all the UEs that do not meet URLLC requirements. While in single connectivity simulations, there are 62 UEs exceeding the 1ms requirement, for the dual connectivity simulations, only 60 of them are. The interesting thing to observe here is that there are actually 13 UEs reducing their delay, but then there are 11 UEs worsening it. From these 13 UEs improving their delay, only 7 of them are in DC. This leads to the conclusion that DC can be useful but the configuration algorithm is very sensitive to the UE and channel conditions. Therefore, it can be useful for some of the UEs but have totally the opposite affect for others.

Was not > 1ms in SC simulations but it is with DC simulations

Is not > 1ms anymore with DC simulations

Was > 1ms for both simulations

Table 7. Per UE Statistics

MT-ID	Dual Connectivity	Serving Cell	BLER Net-0	BLER Net-1	Path Loss to Serving
5	NO	SC	0	0,08148	77,246
6	NO	SC	0	0,0833	81,28
45	NO	SC	0	0,077419	76,81
57	NO	MC	0,0923	0	96,666
68	YES	SC	0,4606	0,0937	94,19
93	NO	SC	0	0,09933	75,22
141	NO	SC	0	0,05737	81,34
142	YES	SC	0,09183	0,1296	91,48
143	YES	SC	0,09401	0,1013	99,53
148	NO	SC	0	0,1086	73,07
150	YES	SC	0,0673	0,125	82,58
162	YES	SC	0,11458	0,1363	106,61
170	YES	SC	0.01235	0,07017	69,32
177	NO	SC	0	0,0697	78,86
196	YES	SC	0,09649	0,1111	87,02
220	YES	SC	0,08035	0,119	90,244
221	NO	SC	0	0,1031	71,82
232	NO	MC	0,08	0	91,75
238	NO	SC	0	0,0909	74,86
241	NO	SC	0	0,0559	84,40
251	YES	SC	0,0769	0,0243	78,58
274	NO	MC	0,0793	0	94,42
296	YES	SC	0,09615	0,06086	76,58
348	NO	SC	0	0,09459	70,31
353	NO	MC	0,0952	0	95,39

363	NO	SC	0	0,1156	74,98
364	NO	SC	0	0,1062	73,20
397	NO	SC	0	0,0775	75,942
398	NO	MC	0,07630	0	110,62
399	NO	MC	0	0,09345	74,195
412	NO	SC	0	0,1153	81,89
415	NO	MC	0,09352	0	102,19
419	YES	SC	0,04255	0,1111	85,89
424	NO	SC	0	0,08219	72,08
435	YES	SC	0,0769	0,1083	85,202
439	YES	SC	0,0865	0,1157	90,74
443	YES	SC	0,0909	0,1339	90,28
447	NO	SC	0	0,13157	72,72
448	YES	SC	0,082644	0,1127	85,72
449	YES	SC	0,01041	0,09821	80,72
464	NO	SC	0	0,1015	74,19
467	NO	MC	0,088	0	104,51
481	YES	SC	0,1008	0,1063	82,77
482	NO	SC	0	0,1037	76,71
484	NO	SC	0	0,0828	70,94
530	YES	SC	0,1009	0,1056	89,46
538	YES	SC	0,088	0,1223	83,06
545	NO	SC	0	0,0731	85,002
559	NO	SC	0	0,0655	78,008
565	YES	SC	0,0857	0,1045	96,001
568	NO	SC	0	0,0687	64,498
573	NO	SC	0	0,0977	85,95
592	YES	SC	0,0891	0,1181	84,69
596	NO	SC	0	0,0909	73,56
598	NO	SC	0	0,0857	73,23
602	NO	SC	0	0,0869	82,20
610	YES	SC	0,09677	0,0849	90,911
618	NO	SC	0	0,0857	74,24
633	NO	SC	0	0,0714	94,59
642	NO	SC	0	0,1209	69,28
643	NO	SC	0	0,0714	73,46
651	YES	SC	0,0695	0,111	77,79
665	NO	SC	0	0,125	80,23
683	NO	SC	0	0,0921	78,95
689	NO	SC	0	0,0948	72,92
712	NO	MC	0,0625	0	85,883
726	YES	SC	0,0923	0,1126	87,61
727	YES	SC	0,0654	0,137	82,87
751	YES	SC	0,01694	0,1397	72,40
779	NO	SC	0	0,0894	83,902
780	NO	MC	0,1007	0	82,94
794	NO	SC	0	0,1023	79,55
795	YES	SC	0,1025	0,0915	84,42

Annex C

Effects of decreasing CRE – Referred from Section 4.2.

Fig. 46 shows SC results for different URLLC loads when CRE is equal to 15dB, while Fig. 47 shows per layer statistics of the same case. If these are compared with Figs. 48 and 49, which show combined and per layer delay statistics, respectively, for the case where CRE is set to 12dB; it can clearly be seen how the bottleneck appears in the macro cells. This bottleneck makes more appreciable the difference in delay performance for the different loads.

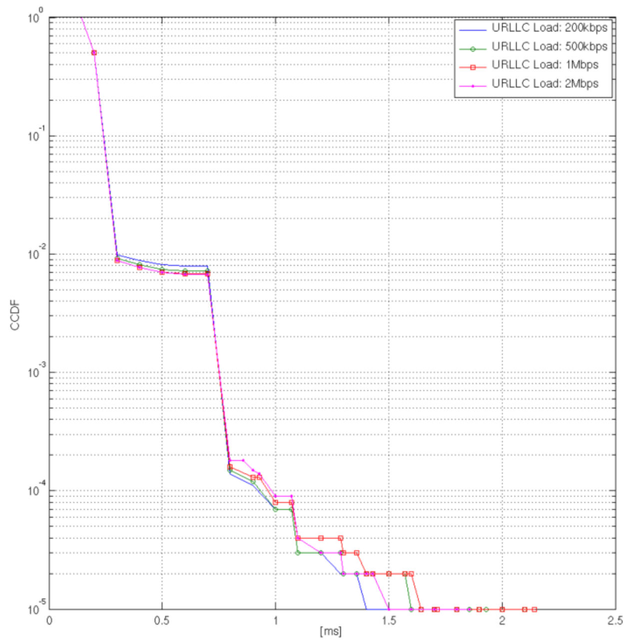

Figure 46. SC Delay Performance for different URLLC loads when CRE = 15dB.

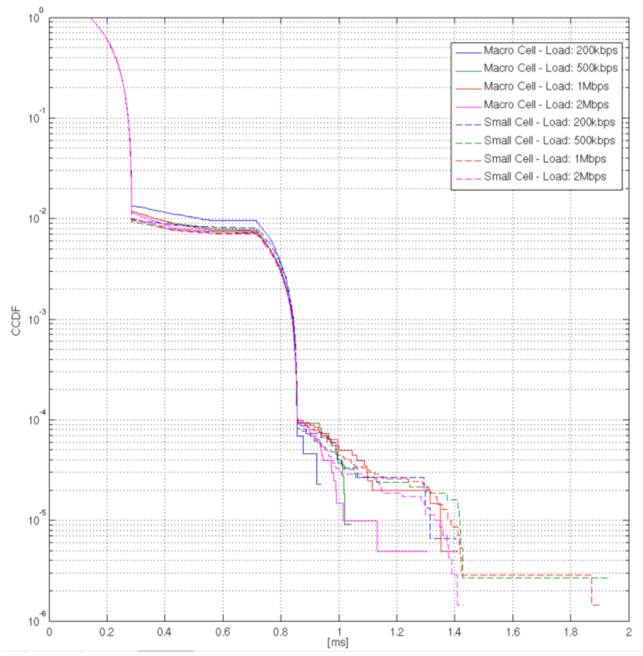

Figure 47. Per Layer SC Delay Performance for different URLLC loads when CRE = 15dB.

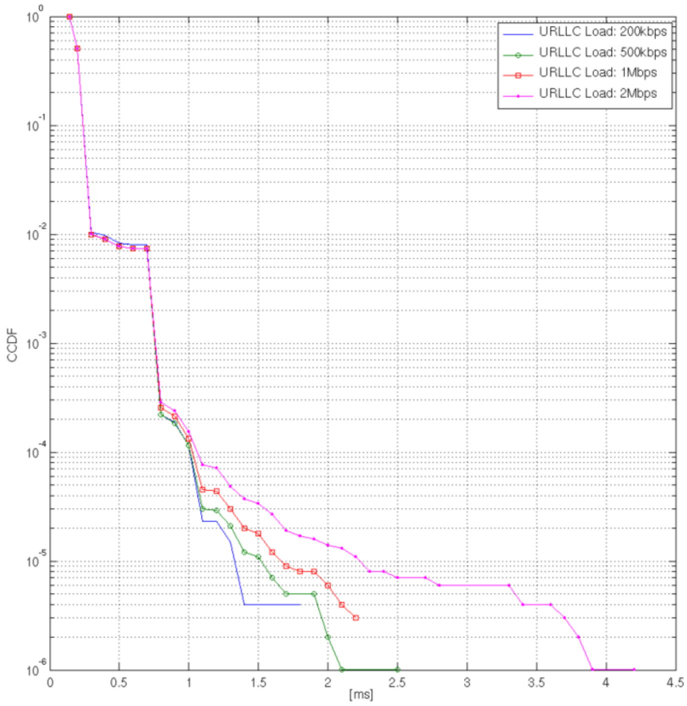

Figure 48.SC Delay Performance for different URLLC loads when CRE = 12dB.

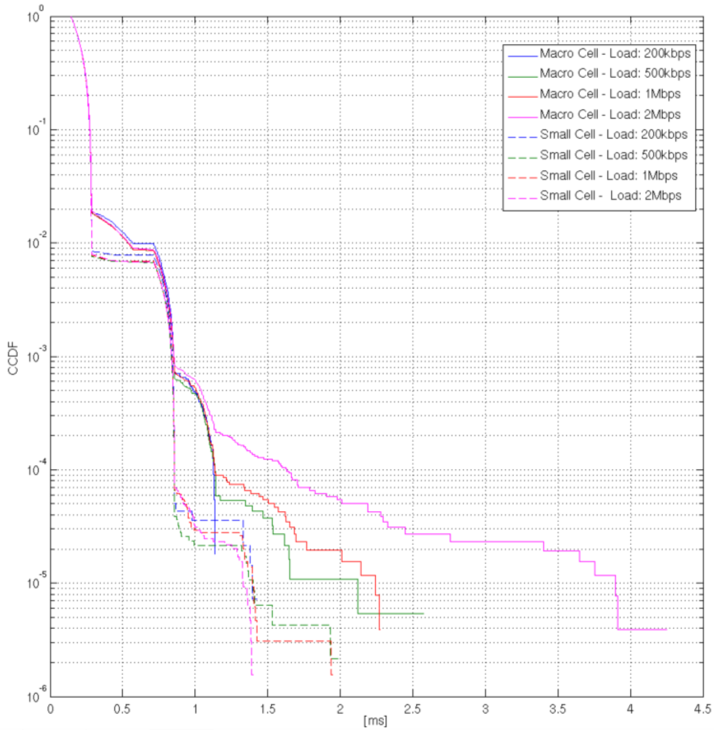

Figure 49. Per Layer SC Delay Performance for different URLLC loads when CRE = 12dB.

BLER Target Performance Comparison – Referred from Section 4.2.

In Fig. 50 can be clearly appreciated how 10% BLER target simulations provide worse results, being far away from meeting URLLC requirements.

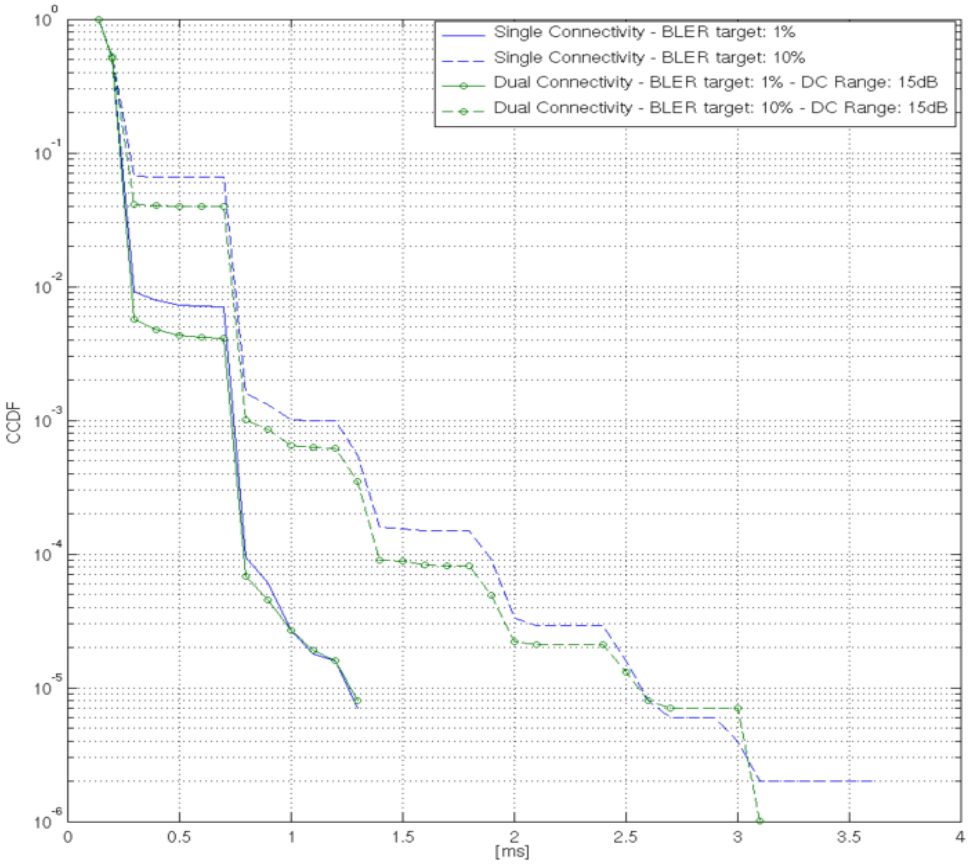

Figure 50. SC vs. DC Delay Performance for different BLER Targets.